



Національний університет «Запорізька політехніка»

## **Радіоелектроніка, інформатика, управління**

Науковий журнал

Виходить чотири рази на рік

№ 1(64) 2023

Заснований у січні 1999 року.

Засновник і видавець – Національний університет «Запорізька політехніка».

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

Запоріжжя

НУ «Запорізька політехніка»

2023

National University «Zaporizhzhia Polytechnic»

## **Radio Electronics, Computer Science, Control**

The scientific journal

Published four times per year

№ 1(64) 2023

Founded in January 1999.

Founder and publisher – National University «Zaporizhzhia Polytechnic».

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

Zaporizhzhia

NU «Zaporizhzhia Polytechnic»

2023

Национальный университет «Запорожская политехника»

## **Радиоэлектроника, информатика, управление**

Научный журнал

Выходит четыре раза в год

№ 1(64) 2023

Основан в январе 1999 года.

Основатель и издатель – Национальный университет «Запорожская политехника».

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

Запорожье

НУ «Запорожская политехника»

2023

**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 ліцензіїю 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;

**Бодянский 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 № 3 від 28.11.2022.**

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](mailto: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 Олеся 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 Ішиномакі 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** – доктор технічних наук, доцент, професор 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. Іллі 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** – Dr. Sc., Associate Professor, 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. Zhukovsky 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 Ursufiu** – 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 № 3 dated 28.11.2022

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](mailto:rvv@zntu.edu.ua)

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

© National University «Zaporizhzhia Polytechnic», 2023

# ЗМІСТ

<b>РАДІОЕЛЕКТРОНІКА ТА ТЕЛЕКОМУНІКАЦІЇ.....</b>	<b>6</b>
<i>Zubkov A. M., Kosovtsov Y. M., Shcherba A. A., Petliuk I. V., Yunda V. A.</i> METHOD OF SELF-DEFENSE OF GROUND (SURFACE) OBJECTS FROM HIGH-PRECISION RADAR MEANS OF AIR SURVEILLANCE.....	6
<b>МАТЕМАТИЧНЕ ТА КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ.....</b>	<b>14</b>
<i>Parfenenko Yu. V., Shendryk V. V., Kholiavka Ye. P., Pavlenko P. M.</i> COMPARISON OF SHORT-TERM FORECASTING METHODS OF ELECTRICITY CONSUMPTION IN MICROGRIDS.....	14
<b>НЕЙРОІНФОРМАТИКА ТА ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ.....</b>	<b>24</b>
<i>Bakumenko N. S., Strilets V. Y., Ugryumov M. L., Zelenskyi R. O., Ugryumova K. M., Starenkiy V. P., Artukh S. V., Nasonova A. M.</i> COMPUTATIONAL INTELLIGENCE METHODS TO PATIENTS STRATIFICATION IN THE MEDICAL MONITORING SYSTEMS.....	24
<i>Lovkin V. M., Subbotin S. A., Oliinyk A. O., Myronenko N. V.</i> METHOD AND SOFTWARE COMPONENT MODEL FOR SKIN DISEASE DIAGNOSIS.....	40
<i>Polyakova M. V.</i> IMAGE SEGMENTATION WITH A CONVOLUTIONAL NEURAL NETWORK WITHOUT POOLING LAYERS IN DERMATOLOGICAL DISEASE DIAGNOSTICS SYSTEMS.....	51
<i>Shelehov I. V., Prylepa D. V., Khibovska Yu. O., Otroshchenko M. S.</i> MACHINE LEARNING DECISION SUPPORT SYSTEMS FOR ADAPTATION OF EDUCATIONAL CONTENT TO THE LABOR MARKET REQUIREMENTS.....	62
<i>Шафроненко А. Ю., Бодяньський Є. В., Головін О. О.</i> КЛАСТЕРИЗАЦІЯ МАСИВІВ ДАНИХ НА ОСНОВІ МОДИФІКОВАНОГО АЛГОРИТМУ СІРОГО ВОВКА.....	73
<b>ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ.....</b>	<b>80</b>
<i>Doroshenko A. Yu., Achour I. Z., Yatsenko O. A.</i> PARAMETER-DRIVEN GENERATION OF EVALUATION PROGRAM FOR A NEUROEVOLUTION ALGORITHM ON A BINARY MULTIPLEXER EXAMPLE.....	80
<i>Kalmykov V. G., Sharypanov A. V., Vishnevskiy V. V.</i> THE CURVE ARC, AS A STRUCTURE ELEMENT OF AN OBJECT CONTOUR IN THE IMAGE TO BE RECOGNIZED.....	89
<i>Olevskiy V. I., Hnatushenko V. V., Korotenko G. M., Olevska Yu. B., Obydennyi Ye. O.</i> APPLICATION OF TWO-DIMENSIONAL PADÉ-TYPE APPROXIMATIONS FOR IMAGE PROCESSING.....	99
<i>Pilkevych I. A., Vakaliuk T. A., Boichenko O. S.</i> THE METHOD OF ASSESSING THE VALUE OF INFORMATION.....	107
<i>Холодна Н. М., Висоцька В. А.</i> ТЕХНОЛОГІЯ ВИПРАВЛЕННЯ ГРАМАТИЧНИХ ПОМИЛОК В УКРАЇНОМОВНОМУ ТЕКСТОВОМУ КОНТЕНТІ НА ОСНОВІ МЕТОДІВ МАШИННОГО НАВЧАННЯ.....	114
<b>УПРАВЛІННЯ У ТЕХНІЧНИХ СИСТЕМАХ.....</b>	<b>141</b>
<i>Batsamut V. M., Hodlevsky S. O.</i> METHOD OF ROUTING A GROUP OF MOBILE ROBOTS IN A FIXED NETWORK FOR SEARCHING THE MISSING OBJECTS IN A TECHNOLOGICAL DISASTER ZONE.....	141
<i>Sabat V. I., Sikora L. S., Durnyak B. V., Povkhan I. F., Polishchuk V. V.</i> ANALYSIS OF RISK TERMINAL FLOWS IN TECHNOGENIC SYSTEMS ARISING IN THE PROCESS OF THREAT IMPACT.....	155

# CONTENTS

<b>RADIO ELECTRONICS AND TELECOMMUNICATIONS.....</b>	<b>6</b>
<i>Zubkov A. M., Kosovtsov Y. M., Shcherba A. A., Petliuk I. V., Yunda V. A.</i> METHOD OF SELF-DEFENSE OF GROUND (SURFACE) OBJECTS FROM HIGH-PRECISION RADAR MEANS OF AIR SURVEILLANCE.....	6
<b>MATHEMATICAL AND COMPUTER MODELING.....</b>	<b>14</b>
<i>Parfenenko Yu. V., Shendryk V. V., Kholiavka Ye. P., Pavlenko P. M.</i> COMPARISON OF SHORT-TERM FORECASTING METHODS OF ELECTRICITY CONSUMPTION IN MICROGRIDS.....	14
<b>NEUROINFORMATICS AND INTELLIGENT SYSTEMS.....</b>	<b>24</b>
<i>Bakumenko N. S., Strilets V. Y., Ugryumov M. L., Zelenskyi R. O., Ugryumova K. M., Starenkiy V. P., Artiukh S. V., Nasonova A. M.</i> COMPUTATIONAL INTELLIGENCE METHODS TO PATIENTS STRATIFICATION IN THE MEDICAL MONITORING SYSTEMS.....	24
<i>Lovkin V. M., Subbotin S. A., Oliinyk A. O., Myronenko N. V.</i> METHOD AND SOFTWARE COMPONENT MODEL FOR SKIN DISEASE DIAGNOSIS.....	40
<i>Polyakova M. V.</i> IMAGE SEGMENTATION WITH A CONVOLUTIONAL NEURAL NETWORK WITHOUT POOLING LAYERS IN DERMATOLOGICAL DISEASE DIAGNOSTICS SYSTEMS.....	51
<i>Shelehov I. V., Prylepa D. V., Khibovska Yu. O., Otroshchenko M. S.</i> MACHINE LEARNING DECISION SUPPORT SYSTEMS FOR ADAPTATION OF EDUCATIONAL CONTENT TO THE LABOR MARKET REQUIREMENTS.....	62
<i>Shafronenko A. Yu., Bodyanskiy Ye. V., Holovin O. O.</i> CLUSTERIZATION OF DATA ARRAYS BASED ON THE MODIFIED GRAY WOLF ALGORITHM.....	73
<b>PROGRESSIVE INFORMATION TECHNOLOGIES.....</b>	<b>80</b>
<i>Doroshenko A. Yu., Achour I. Z., Yatsenko O. A.</i> PARAMETER-DRIVEN GENERATION OF EVALUATION PROGRAM FOR A NEUROEVOLUTION ALGORITHM ON A BINARY MULTIPLEXER EXAMPLE.....	80
<i>Kalmykov V. G., Sharypanov A. V., Vishnevskiy V. V.</i> THE CURVE ARC, AS A STRUCTURE ELEMENT OF AN OBJECT CONTOUR IN THE IMAGE TO BE RECOGNIZED.....	89
<i>Olevskiy V. I., Hnatushenko V. V., Korotenko G. M., Olevska Yu. B., Obydennyi Ye. O.</i> APPLICATION OF TWO-DIMENSIONAL PADÉ-TYPE APPROXIMATIONS FOR IMAGE PROCESSING.....	99
<i>Pilkevych I. A., Vakaliuk T. A., Boichenko O. S.</i> THE METHOD OF ASSESSING THE VALUE OF INFORMATION.....	107
<i>Kholodna N., Vysotska V.</i> TECHNOLOGY FOR GRAMMATICAL ERRORS CORRECTION IN UKRAINIAN TEXT CONTENT BASED ON MACHINE LEARNING METHODS.....	114
<b>CONTROL IN TECHNICAL SYSTEMS.....</b>	<b>141</b>
<i>Batsamut V. M., Hodlevsky S. O.</i> METHOD OF ROUTING A GROUP OF MOBILE ROBOTS IN A FIXED NETWORK FOR SEARCHING THE MISSING OBJECTS IN A TECHNOLOGICAL DISASTER ZONE.....	141
<i>Sabat V. I., Sikora L. S., Durnyak B. V., Povkhan I. F., Polishchuk V. V.</i> ANALYSIS OF RISK TERMINAL FLOWS IN TECHNOGENIC SYSTEMS ARISING IN THE PROCESS OF THREAT IMPACT.....	155

# РАДІОЕЛЕКТРОНІКА ТА ТЕЛЕКОМУНІКАЦІЇ

## RADIO ELECTRONICS AND TELECOMMUNICATIONS

UDC 621.396.96

### METHOD OF SELF-DEFENSE OF GROUND (SURFACE) OBJECTS FROM HIGH-PRECISION RADAR MEANS OF AIR SURVEILLANCE

**Zubkov A. M.** – Dr. Sc., Senior Research Associate, Leading Researcher of Army Scientific Center Research Department (Missile Troops and Artillery) of the Scientific Center of the Land Forces, Hetman Petro Sahaidachnyi National Army Academy, Lviv, Ukraine.

**Kosovtsov Y. M.** – PhD, Senior Research Associate, Leading Researcher of Army Scientific Center Research Department (Troops Training) of the Scientific Center of the Land Forces, Hetman Petro Sahaidachnyi National Army Academy, Lviv, Ukraine.

**Shcherba A. A.** – PhD, Associate Professor, Deputy Chief of Artillery Reconnaissance Complexes and Devices Department, Hetman Petro Sahaidachnyi National Army Academy, Lviv, Ukraine.

**Petliuk I. V.** – PhD, Leading researcher of the research laboratory (information and geo-information systems) of the research department (military management systems) of the Scientific Center of the Land Forces, Hetman Petro Sahaidachnyi National Army Academy, Lviv, Ukraine.

**Yunda V. A.** – PhD, Associate Professor, Head of Missile Troops Department, Hetman Petro Sahaidachnyi National Army Academy, Lviv, Ukraine.

#### ABSTRACT

**Context** it is caused by the need to search for scientific and technical ways to ensure the effectiveness of protecting ground (surface) objects from high-precision guided missile weapons.

**Objective** it is a necessity to ensure effective self-defense of objects from radar homing means.

**Method.** Electrodynamics modeling of Echo signals from spatially distributed objects, taking into account the features of their design and related operational limitations.

**Results.** Based on the analysis of the shortcomings of the well-known method of protecting stationary objects from radar surveillance and damage, based on the simulation of an effective reflection center outside the physical dimensions of the object, a new method of countering high-precision measurement of coordinates of stationary and mobile ground (surface) objects is proposed. The technique is based on the spatial deformation of the location of the effective target reflection center with dynamics that exceed the inertial capabilities of the auto-observation contour of the attacking missile (projectile). A structural and functional scheme of technical implementation of the methodology based on the first proposed relationship of simple design and technological solutions is proposed and justified.

**Conclusions.** The analytical model of Echo signals of spatially distributed ground (surface) objects was further developed, which takes into account the specifics of their design, and on its basis, for the first time, a universal method of self-defense of objects from radar home-leading devices was developed, which is implemented in a patented method and complex to exclude damage to protected objects.

**KEYWORDS:** radar home-leading head, electromagnetic wave scattering model, local reflection section, direction finding characteristic, angular reflector, polarization grating.

#### ABBREVIATIONS

CU is a control unit;

SGRH is a self-guided radar head;

SP is a scatter plot

LDA is a local display area;

CR is a corner reflector;

MC is a motor controller;

RD is a reduction drive;

RS is a radar station;

RCS is a radar cross-section;

SCS is a scattering cross-section

HLH is a home-leading head;

BSD is a Backscattering diagram

LDA is a local display area;

AR is an angle reflector;

RS is a radar set

RO is a radar objects;

ESS is an effective scattering surface;

ESC is an effective scattering center.

## NOMENCLATURE

$\alpha$  is an angle of view of the target by Yaw;

$\varepsilon$  is a Pitch angle of view of the target;

$\vec{\gamma}$  is a vector that characterizes the observation

conditions and object orientation;

$A_i$  is an amplitude of the signal reflected from  $i$  LDA;

$R_i$  is a radius-vector  $i$  LDA;

$E$  is an electromagnetic field strength;

$\omega$  is a circular frequency;

$\lambda$  is a wavelength of the probing signal;

$l_{\alpha,\varepsilon}$  is a geometric size of the target or a fragment of its structure in the “picture” plane;

$\Delta\theta_{\alpha,\varepsilon}$  is a width of the linear section of the direction finding characteristic by Yaw  $\alpha$  and Pitch  $\beta$ ;

$\Delta f_e$  is an effective spectral band of the rocket control circuit;

$\sigma_{\alpha,\varepsilon}$  is an error of auto-tracking of the target by Yaw  $\alpha$  and Pitch  $\beta$ .

## INTRODUCTION

Radar means of air (space) observation is the only effective tool for highly informative remote monitoring of the Earth’s surface in the interests of solving various general technical and special tasks in the absence of optical transparency of the surface layer of the atmosphere.

An urgent scientific and applied problem is minimizing the probability of high-precision weapons hitting ground (surface) equipment objects. There is a well-known method of self-defense of an object by installing a simulator outside it, for example, in the form of an angle reflector.

The disadvantages of this approach include:

– use only for stationary objects, since the AR ESS must be guaranteed to exceed the ESS of the protected object;

– a narrow corner protected area, which leads to an increase in the number of AR in conditions of a priori uncertainty in the direction of attack.

Taking into account the above, there is a scientific and technical task-the development of methodological and instrumental bases for electrodynamic simulation of Echo signals of stationary and moving objects in the radar home-leading channel, which exclude the defeat of the protected object regardless of the direction of attack, and optimal according to the criterion “efficiency/cost”.

Thus, the topic that involves the search for scientific and technical ways to ensure effective protection of ground (surface) objects (targets) from high-precision missile weapons is relevant.

**Object of research** there is a process of forming echo signals from ground (surface) targets in the radar home-leading channel.

**Subject of research** is a analytical model of Echo signals for developing a method of self-defense of targets and a method of its practical implementation.

**Purpose of the work** is a ensuring effective self-defense of objects from radar home-leading means, for which it is necessary:

– perform an analysis of known approaches to describing echoes from spatially distributed targets;

– to develop and substantiate a model of scattering of electromagnetic waves in the radio range from the forming structure of observed objects in the form of a set of LDA and its analytical description;

– based on the LDA model, develop a methodology and method for self-defense of objects from homing means and perform a model experiment to analyze their effectiveness.

## 1 PROBLEM STATEMENT

It is known that the energy characteristics of the echo signal, which determine the maximum range of the homing section and potential accuracy, depend on the target’S EER, and the total guidance error is a function of the dynamics of missile-target movement [1–3]. At the same time, the missile’s radar homing contour tracks the angular position of the target’S ESC [4–7].

The analytical criterion for stable operation of the homing circuit is the ratio

$$\sigma_{\alpha,\varepsilon} \ll \Delta\theta_{\alpha,\varepsilon}. \quad (1)$$

To ensure the failure of auto wiring, artificial provision of the condition is necessary

$$\sigma_{\alpha,\varepsilon} > \Delta\theta_{\alpha,\varepsilon} \quad (2)$$

per hour  $t > \frac{1}{\Delta f_e}$ .

For the first time, a complex is proposed as an instrumental basis for implementing this condition, which includes a set of AR that rotate asynchronously and provide a dynamic stochastic change in the BSD of the protected object.

## 2 REVIEW OF THE LITERATURE

In the radar surveillance channel, two approaches are used to model echoes from spatially distributed targets [8–13].

The phenomenological model is based on direct observations of the scattering process of the probing signal of the forming surface of the target. In practice, two types of phenomenological models are used:

– Radial model, which is the basis of the method of geometric optics and geometric diffraction theory. The radial representation of reflected waves is the main feature of the model. Secondary effects are diffraction and polarization. The model adequately describes the scattering process when the condition is met:

$$l_{\alpha,\varepsilon} > \lambda;$$

– a wave model based on the Huygens-Fresnel principle (physical optics method). The Shape of the object and the angle of the observation point relative to the observed object play a crucial role.

The analog model is based not on direct observations of the process of radio wave scattering, but on the results

of studies of other phenomena that occur in a similar way to the simulated process. At the same time, as in the phenomenological model, the main features of the process are highlighted:

– the “shiny dots” model. It is based on observations of light reflection from polished target layouts. This model is used to analyze the field reflected from rough surfaces;

– facet model. It is based on the approximation of the target surface in the form of a set of flat reflectors that are normally oriented to the incident wave and on the observation of light reflection from the water surface.

The limited possibilities of applying the above models in practice are due to:

– the need to detail a priori information for a specific phono-target situation accompanying the homing of a missile (projectile);

– analytical complexity of obtaining the resulting expressions describing the echo signal from real objects that are observed;

– weak resistance to changes in observation conditions (in particular, the object’s angle).

Therefore, the development of these models in the direction of universality of application while ensuring adequacy should be considered an urgent scientific and applied task.

### 3 MATERIALS AND METHODS

In order to ensure the adequacy of modeling the scattering process of sensing signals to real physical phenomena accompanying observations of ground (surface) objects in the radar channel, a model is proposed that combines the capabilities of the phenomenological and analog models discussed above – the LDA model [14–16].

The essence of the model is based on the following prerequisites:

– the field scattered by a spatially distributed object is formed by a small number of waves, the source of which is located on the “illuminated” part of its forming surface;

– the distances between LDA exceed the wavelength, and the geometric area that they occupy is small compared to the area of the entire “illuminated” part of the object’s surface;

– LDA are partially coherent, but can contain pairs that are either completely coherent or completely incoherent;

– the location of the LDA is clearly related to the design features of the forming surface of the observed object.

Analytically, the field scattered by the LDA aggregate can be represented as

$$E(t, \omega, \vec{\gamma}) = \sum_{i=1}^I A_i(\omega \vec{\gamma}) e^{j\omega \frac{2R_i(\vec{\gamma})}{c}}. \quad (3)$$

The method of self-defense of a ground (surface) object provides for artificial provision of the condition (2). The latter can be achieved by asynchronously changing  $A_i$  in expression (3) due to the rotation of angular reflectors located along the perimeter of the protected object.

This leads to a chaotic change in the effective scattering center of the target with dynamics that exceed the capabilities of the missile (projectile) homing contour.

The practical implementation of the methodology is illustrated in Fig. 1 [17].

Structurally, the complex includes a set of an angular reflector with a polarization grating in the aperture to achieve polarization invariance. Asynchronous rotation of the angle reflector is provided by a controlled electric drive through a gearbox by connecting to the control unit.

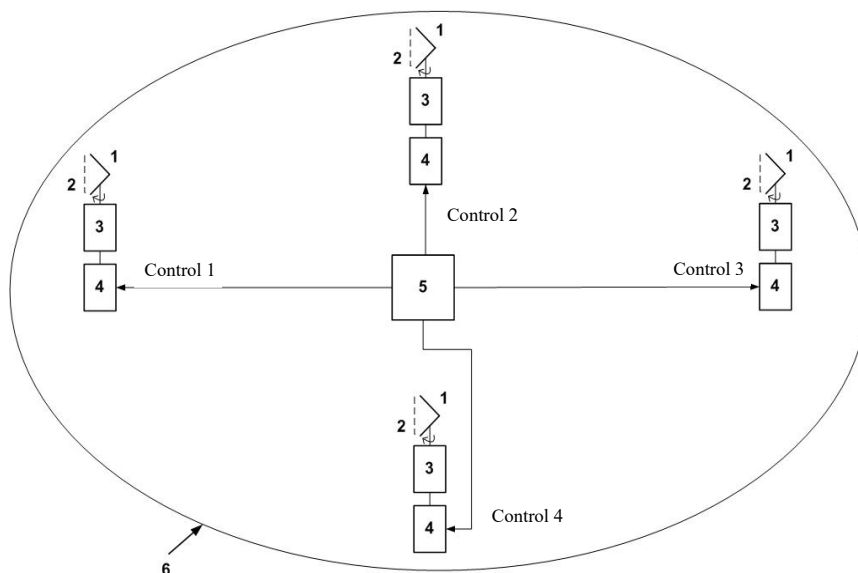


Figure 1 – The structural and functional diagram of a complex for simulating a ground (surface) object in a radar homing channel  
 1 is an angle reflector; 2 is a polarizing grating; 3 is a gearbox; 4 is a controlled electric drive; 5 is a Control Unit;  
 6 is a geometric contour of the protected object



#### 4 EXPERIMENTS

A model experiment was performed [18] to prove the adequacy of the proposed LDA model to the real conditions for applying the developed method of self-defense of ground (surface) objects. The essence of the experiment is electrodynamic modeling of Echo signals of complex RO: a railway bridge, a missile system launcher, a helicopter (priority objects of destruction) in a radar home-leading channel. The inclined range was 10 000 m,  $\varepsilon = 75...85$  angle degrees and  $\alpha = 0...180$  angle degrees,  $\lambda = 8$  and 3 mm, and frequencies of 36 and 95 GHz.

To calculate the ratio of the ESS of complex RO when irradiated with an electromagnetic wave with a probing signal wavelength of 8 mm to the ESS of complex RO when irradiated with an electromagnetic wave with  $\lambda = 3$  mm in the Maple 15 Medium, data arrays were created at Target viewing angles  $\varepsilon = 75...85$  angle degrees and  $\alpha = 0...180$  angle degrees in increments of 1 angle degrees.

Polygonal models of the bridge, launcher, and helicopter are shown in Fig. 2-4.

#### 5 RESULTS

The resulting BSD of these objects at wavelengths of 8 and 3 mm are shown in Fig. 5 – 7, respectively.

The obtained patterns allow for a clear physical interpretation.

A graphical representation of the ratio of the ESS of complex RO when irradiated with an electromagnetic

wave with  $\lambda = 8$  mm to the ESS of complex RO when irradiated with an electromagnetic wave with  $\lambda = 3$  mm in a linear ratio is shown in Fig. 8, respectively.

Plots with a positive value of  $10\lg(\sigma_{\lambda=8\text{mm}}/\sigma_{\lambda=3\text{mm}})$  correspond to the case when the ESS of a complex RO when irradiated with an electromagnetic wave with  $\lambda = 8$  mm is greater than the ESS of the same object when irradiated with an electromagnetic wave with  $\lambda = 3$  mm (fig. 8).

Depending on the observation conditions, the value of this ratio can be either greater than 1 or less. Based on the dependence of the detection probability on the signal/noise ratio at the input of the linear detector, the probability of detection in each of these sections is determined with a false alarm probability of  $10^{-6}$  and the fixed probability of detecting a complex RO as a whole is 0.9.

Under such conditions, the value of the signal/noise ratio at the input of the linear detector will be 15 dB, then in the section of the ESS ratio, where the ESS with  $\lambda = 8$  mm is greater than the ESS with  $\lambda = 3$  mm, the probability varies in the range of 0.9–1.

In the area where the ESS with  $\lambda = 8$  mm is 2 times larger than the ESS with  $\lambda = 3$  mm or less, the signal – noise ratio will be 13–15 DB, and the detection probability will be 0.7–0.9.



Figure 2 – Polygonal bridge model

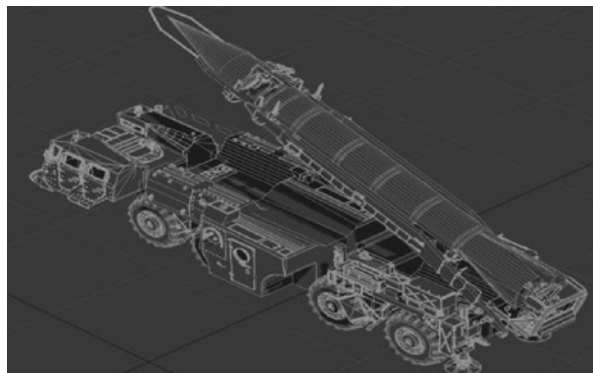


Figure 3 – Polygonal model of a rocket system launcher



Figure 4 – Polygonal helicopter model

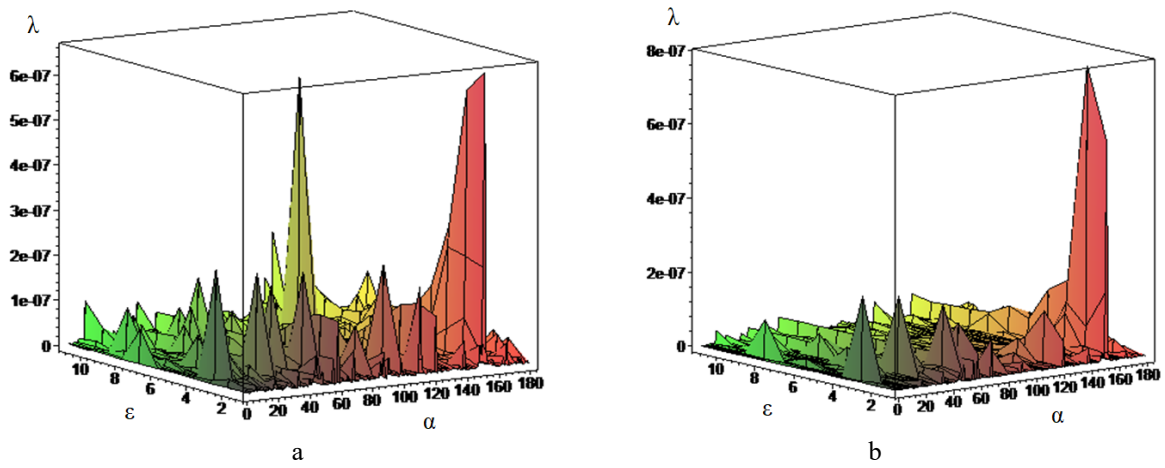


Figure 5 – Bridge backscattering diagram: a –  $\lambda = 8$  mm; b –  $\lambda = 3$  mm;

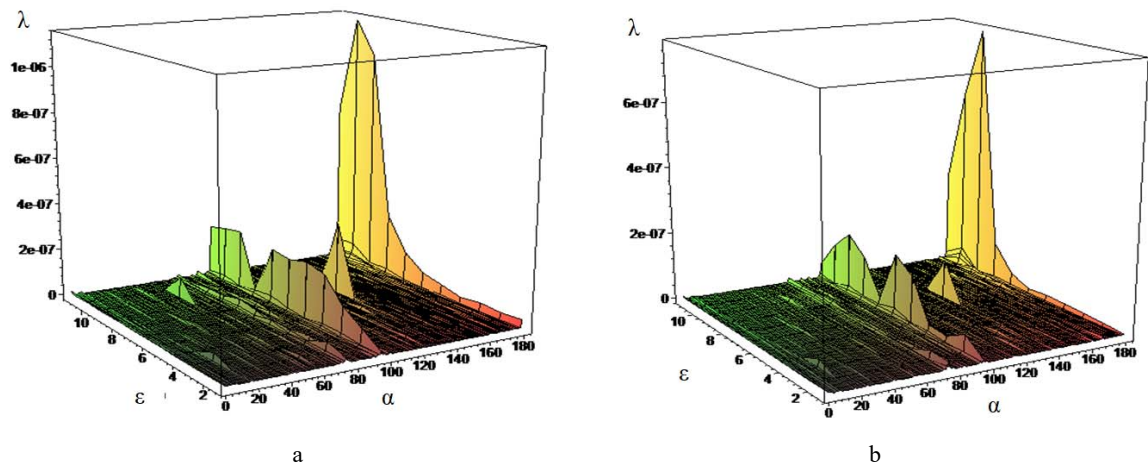


Figure 6 – Backscattering diagram of the launcher: a –  $\lambda = 8$  mm; b –  $\lambda = 3$  mm;

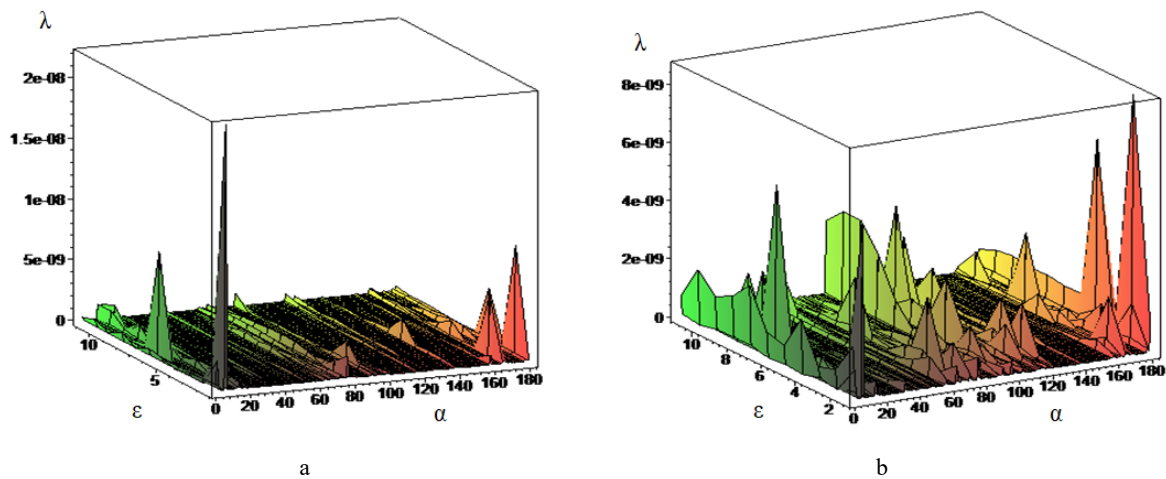


Figure 7 – Helicopter backscattering diagram: a –  $\lambda = 8$  mm; b –  $\lambda = 3$  mm.

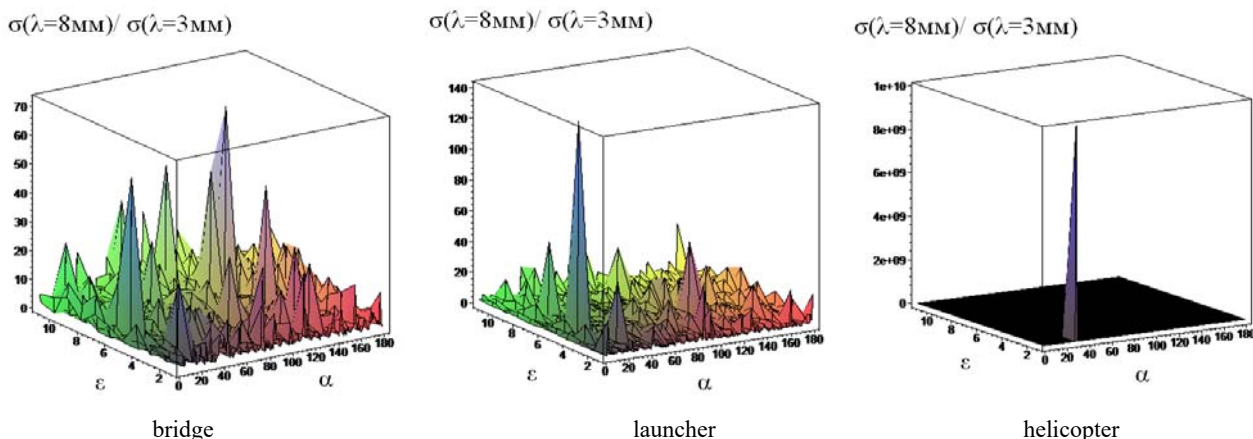


Figure 8 – The ratio of the ESS of complex RO when irradiated with an electromagnetic wave with  $\lambda=8$  mm to the ESS of complex RO when irradiated with an electromagnetic wave with  $\lambda=3$  mm in a linear ratio

As can be seen from the nomogram (Fig. 8), Blue shows the area where the ESS with  $\lambda = 8$  mm is larger, yellow shows the area where the ESS with  $\lambda = 8$  mm is 2 times larger.

The results of the experiment confirm the initial prerequisites, namely:

- the total energy of target Echo signals is determined by a limited number of BSD extremes associated with the spatial distribution of LDA;
- since the HLH “works” on the ESC, an artificial dynamic change in its location is a universal means of self-defense of an object from self-guided means.

## 6 DISCUSSION

The method of the experiment provided for:

- approximation of the forming surface of targets observed in the radar channel in the form of an unbroken set of tangent triangles;
- formation of target Echo signals as a superposition of elementary Echo signals from triangles;
- obtaining backscattering diagrams of observed objects at an inclined range of 10,000 m at Target viewing angles of 75...85 angle. degrees in pitch and 0...180 angle. degrees at yaw.

The developed method, in comparison with the known ones, makes it possible to significantly expand the range of practical application conditions, removes restrictions on the type of protected object, its design, and the presence or absence of movement.

The reliability of the results obtained is confirmed by the possibility of their clear physical interpretation and modeling data. The possibility of wide application of the technique for protecting ground (surface) objects from radar home-leading means is based on the simplicity of technical implementation and low cost in comparison with known approaches

For the first time, an analytical description of the echo signal scattered by a complex object (3) was developed, together with the mandatory fulfillment of Condition (2), which gives an adequate description of the processes that accompany the observation of a spatially distributed

target and is the basis of the developed method of its self-defense.

The practical implementation of the technique, in contrast to the known approaches, can technically be carried out in accordance with the proposed method for a wide range of external conditions and different dynamics of mutual movement of the protected and attacking objects.

It is important to note that the proposed method and method of self-defense are universal, since:

- invariant to the design and material of the forming surface of the protected object;
- effective for any trajectory and number of attacking objects;
- allow electrodynamic modeling to quantify the effectiveness of self-defense.

## CONCLUSIONS

**Has been solved a scientific problem** of methodological support of effective self-defense of ground (surface) objects from radar home-leading combat elements by stochastic change in the effective scattering surface of the protected object with dynamics exceeding the speed of the home-leading contour of attacking elements for what purpose:

- the analysis of the shortcomings of known approaches based on simulating the effective reflection center of the protected object beyond its physical dimensions is performed;
- for the first time, a universal model of a spatially distributed target observed in a radar home-leading channel in the form of a limited set of LDA is proposed and justified;
- based on the new model of a spatially distributed target, a system of self-defense of ground (surface) objects from radar homing means is proposed for the first time. It is based on the deformation of the location of the effective reflection center with dynamics that exceed the inertial capabilities of the home-leading contour of attacking means.

**Practical significance** the obtained results are determined by the proposed method and complex of self-defense of ground (surface) objects, the priority of which is confirmed by the patent for the invention, as well as the data of the model experiment.

**Directions for further research** there is an optimization of the number of AR based on the “efficiency/cost” criterion with reference to the volume (area) of the protected object.

#### REFERENCES

1. Kurkotkin V. I., Sterligov V. L. Home-leading missiles. Moscow, Voenizdat, 1990, 226 p. (in Russian).
2. Kingsley S., Quegan Sh. Understanding Radar Systems. Mendham NJ, Sci Tech Publishing, 1999, 375 p. DOI: 10.1049/SBRA034E
3. Chen R. H., Speyer J. L., Lianos D. Optimality of Error Dynamics in Missile Guidance Problems, *Journal of Guidance, Control, and Dynamics*, 2012, Vol. 30, № 6, pp. 1579–1589. DOI: 10.2514/1.30107
4. He Sh., Lee Ch.-H. Homing Missile Guidance and Estimation Under Agile Target Acceleration, *Journal of Guidance, Control, and Dynamics*, 2018, Vol. 41, № 7, pp. 1624–1633. DOI: 10.2514/1.G003343
5. Boarov S. Evaluation accuracy for homing system of ground-air missile, *Vojnotehnicki Glasnik*, 2006, Vol. 54, № 2, pp. 151–159. DOI: 10.5937/vojtehg0602151b
6. De Maio A., Greco M. S. Modern Radar Detection Theory. Edison, NJ, 2015, 400 p. DOI: 10.1049/SBRA509E
7. Miwa S., Imado F. Clutter effect on the guidance of a semi-active radar homing missile, *Journal of Guidance, Control, and Dynamics*, 1986, Vol. 9, № 3, pp. 268–273. DOI: 10.2514/3.20102
8. Ostrovityanov R. V., Basalov F. A. statistical theory of long-range target radar. Moscow, Radio i svyaz, 1982, 232 p. (in Russian).
9. Shtager E. A. Scattering of radio waves on complex bodies. Moscow, Radio i svyaz, 1986, 184 p. (in Russian).
10. Khrychov V. S., Legenkiy M. N. About modeling the waves scattering on the complex shape objects, *Visnyk of V.N. Karazin Kharkiv National University, series “Radio Physics and Electronics”*, 2018, Vol. 29, pp. 50–56. DOI: 10.26565/2311-0872-2018-29-07
11. Sukharevsky O. I. Electromagnetic Wave Scattering by Aerial and Ground Radar Objects. Boca Raton, CRC Press, 2015, 334 p. DOI: 10.1201/9781315214511
12. Sun G., Wang J., Qin Sh., Na J. Radar target recognition based on the multi-resolution analysis theory and neural network, *Pattern Recognition Letters*, 2008, Vol. 29, № 16, pp. 2109–2115. DOI: 10.1016/j.patrec.2008.07.006
13. Luo, S., Li, S. Automatic target recognition of radar HRRP based on high order central moments features, *Journal of Electronics (China)*, 2009, Vol. 26, pp. 184–190. DOI: 10.1007/s11767-007-0111-3
14. Legenkiy M., Khrychov V. Numerical modeling of electromagnetic scattering from complex shape object with coating, *Frequenz*, 2022, Vol. 76, №. 1–2. pp. 75–82. DOI: 10.1515/freq-2021-0062
15. Bass F. G., Freilikher V. D., Presentsov V. V. Electromagnetic Wave Scattering From Small Scatterers of Arbitrary Shape, *Journal of Electromagnetic Waves and Applications*, 2000, Vol. 14, № 3, pp. 269–283. DOI: 10.1163/156939300X00789
16. Zubkov A. M., Krasnik Ya. V., Martynenko S. A., Yunda V. A., Andreev I. M. Method of self-defense of ground objects from high-precision radar means of air surveillance. *Modern problems and achievements in the field of radio engineering, telecommunications and information technologies : X International conference, Zaporizhzhia, 7–9 October 2020, proceedings*, NU “Zaporizhzhia Polytechnic”, 2020, pp. 23–24.
17. Method of self-defense of ground or surface objects from radar homing devices and a complex for its implementation: PJSC. 122340 Ukraine: IPC (2020.01) F41H 11/02 (2006.01), F41J 2/00. no. a201800937; application form. 01.02.2018; publ. 26.10.2020, Bul. № 20.
18. Yunda V. A., Zubkov A. M., Kosovtsov YU. M., Atamanyuk V. V., Mocherad V. S. Komp'yuterne modelyuvannya mul'tyspektral'noyi tsifonovoyi obstanovky pry samonavedenni rakety na nazemnu tsil', *Nauka i tekhnika Povitrynykh Syl Zbroynykh Syl Ukrainy*, Vol. 2016, № 1, pp. 151–155.

Received 12.01.2023.

Accepted 03.02.2023.

УДК 621.396.96

#### МЕТОДИКА САМОЗАХИСТУ НАЗЕМНИХ (НАДВОДНИХ) ОБ'ЄКТІВ ВІД ВИСОКОТОЧНИХ РАДІОЛОКАЦІЙНИХ ЗАСОБІВ ПОВІТРЯНОГО СПОСТЕРЕЖЕННЯ

**Зубков А. М.** – д-р техн. наук, с.н.с., провідний науковий співробітник науково-дослідного відділу (ракетних військ та артилерії) Наукового центру Сухопутних військ Національної академії сухопутних військ імені гетьмана Петра Сагайдачного, Львів, Україна.

**Косовцов Ю. М.** – канд. фіз.-мат. наук, с.н.с., провідний науковий співробітник науково-дослідного відділу (ракетних військ та артилерії) Наукового центру Сухопутних військ Національної академії сухопутних військ імені гетьмана Петра Сагайдачного, Львів, Україна.

**Щерба А. А.** – канд. техн. наук, доцент, заступник начальника кафедри комплексів та приладів артилерійської розвідки факультету ракетних військ і артилерії Національної академії сухопутних військ імені гетьмана Петра Сагайдачного, Львів, Україна.

**Петлюк І. В.** – канд. техн. наук, провідний науковий співробітник науково-дослідної лабораторії (інформаційних та геоінформаційних систем) науково-дослідного відділу (систем управління військами) Наукового центру Сухопутних військ Національної академії сухопутних військ імені гетьмана Петра Сагайдачного, Львів, Україна.

**Юнда В. А.** – канд. техн. наук, доцент, начальник кафедри ракетних військ факультету ракетних військ і артилерії Національної академії сухопутних військ імені гетьмана Петра Сагайдачного, Львів, Україна.

## АНОТАЦІЯ

**Актуальність** зумовлена необхідністю пошуку науково-технічних шляхів забезпечення ефективності захисту наземних (надводних) об'єктів від високоточного ракетного керованого озброєння.

**Метою роботи** є забезпечення ефективного самозахисту об'єктів від радіолокаційних самонавідних засобів.

**Метод.** Електродинамічне моделювання ехо-сигналів від просторово розподілених об'єктів з урахуванням особливостей їх конструкції та супутніх експлуатаційних обмежень.

**Результати.** На основі аналізу недоліків відомої методики захисту стаціонарних об'єктів від радіолокаційних засобів спостереження і ураження, оснований на імітації ефективного центру відображення за межами фізичних розмірів об'єкту, запропонована нова методика протидії високоточному вимірюванню координат стаціонарних і рухомих наземних (надводних) об'єктів. Методика базується на просторовій деформації місцезнаходження ефективного центру відбиття цілі з динамікою, яка перевищує інерційні можливості контуру авто спостереження атакуючої ракети (снаряду). Запропонована і обґрунтована структурно-функціональна схема технічної реалізації методики на основі вперше запропонованого взаємозв'язку нескладних конструкторсько-технологічних рішень.

**Висновки.** Отримала подальший розвиток аналітична модель ехо-сигналів просторово розподілених наземних (надводних) об'єктів, яка враховує специфіку їх конструкції, та на її основі вперше розроблена універсальна методика самозахисту об'єктів від радіолокаційних засобів самонаведення, яка реалізована у запатентованому способі та комплексі для виключення ураження захищасмих об'єктів.

**КЛЮЧОВІ СЛОВА:** радіолокаційна головка самонаведення, модель розсіювання електромагнітних хвиль, ділянка локального відображення, пеленгаційна характеристика, кутовий відбивач, поляризаційна решітка.

## ЛІТЕРАТУРА

1. Куркоткин В. И. Самонаведение ракет / В. И. Куркоткин, В. Л. Стерлигов. – Москва : Воениздат, 1990. – 226 с.
2. Kingsley S. Understanding Radar Systems / S. Kingsley, Sh. Quegan. – Mendham NJ : Sci Tech Publishing, 1999. – 375 p. DOI: 10.1049/SBRA034E
3. Chen R. H. Optimality of Error Dynamics in Missile Guidance Problems / R. H. Chen, J. L. Speyer, D. Lianos // Journal of Guidance, Control, and Dynamics. – 2012. – Vol. 30, № 6. – P. 1579–1589. DOI: 10.2514/1.30107
4. He Sh. Homing Missile Guidance and Estimation under Agile Target Acceleration / Sh. He, Ch.-H. Lee // Journal of Guidance, Control, and Dynamics. – 2018. – Vol. 41, № 7. – P. 1624–1633. DOI: 10.2514/1.G003343
5. Boarov S. Evaluation accuracy for homing system of ground-air missile / S. Boarov // Vojnotehnicki Glasnik. – 2006. – Vol. 54, № 2. – P. 151–159. DOI: 10.5937/vojtehg0602151b
6. De Maio A. Modern Radar Detection Theory / A. De Maio, M. S. Greco. – Edison, NJ : Scitech Publishing, 2015. – 400 p. DOI: 10.1049/SBRA509E
7. Miwa S. Clutter effect on the guidance of a semi-active radar homing missile / S. Miwa, F. Imado // Journal of Guidance, Control, and Dynamics. – 1986. – Vol. 9, № 3. – P. 268–273. DOI: 10.2514/3.20102
8. Островитянов Р. В. Статистическая теория радиолокации протяжённых целей / Р. В. Островитянов, Ф. А. Басалов. – Москва : Радио и связь, 1982. – 232 с.
9. Штагер Е. А. Рассеяние радиоволн на телах сложной формы / Е. А. Штагер. – Москва : Радио и связь, 1986. – 184 с.
10. Хричов В. С. Щодо моделювання розсіювання хвиль на об'єкті складної форми / В. С. Хричов, М. М. Легенький // Вісник Харківського національного університету імені В. Н. Каразіна. Серія «Радіофізика та електроніка». – 2018. – № 29. – С. 50–56. DOI: 10.26565/2311-0872-2018-29-07
11. Sukharevsky O. I. Electromagnetic Wave Scattering by Aerial and Ground Radar Objects (1st ed.) / O. I. Sukharevsky. – Boca Raton : CRC Press, 2015. – 334 p. DOI: 10.1201/9781315214511
12. Radar target recognition based on the multi-resolution analysis theory and neural network / [G. Sun, J. Wang, S. Qin, J. Na] // Pattern Recognition Letters. – 2008. – Vol. 29, № 16. – P. 2109–2115. DOI: 10.1016/j.patrec.2008.07.006
13. Luo S. Automatic target recognition of radar HRRP based on high order central moments features / S. Luo, S. Li // Journal of Electronics (China). – 2009. – Volume 26. – P. 184–190. DOI: 10.1007/S11767-007-0111-3
14. Legenkiy M. Numerical modeling of electromagnetic scattering from complex shape object with coating / M. Legenkiy, V. Khrychov // Frequenz. – 2022. – Vol. 76, № 1–2. P. 75–82. DOI: 10.1515/freq-2021-0062
15. Bass F. G. Electromagnetic Wave Scattering From Small Scatterers of Arbitrary Shape / F. G. Bass, V. D. Freilikher, V. V. Prosentsov // Journal of Electromagnetic Waves and Applications. – 2000. – Vol. 14, № 3. – P. 269–283, DOI: 10.1163/156939300X00789
16. Методика самозахисту наземних об'єктів від високоточних радіолокаційних засобів повітряного спостереження / [А. М. Зубков, Я. В. Красник, С. А. Мартиненко та ін.] // Сучасні проблеми і досягнення в галузі радіотехніки, телекомунікацій та інформаційних технологій : X Міжнародна науково-практична конференція, Запоріжжя, 7–9 жовтня 2020 р. : тези доповідей. – Запоріжжя: НУ «Запорізька політехніка», 2020. – С. 23–24.
17. Пат. 122340 Україна, МПК 2020.01 F41H 11/02 (2006.01), F41J 2/00. Спосіб самозахисту наземних або надводних об'єктів від радіолокаційних засобів самонаведення та комплекс для його реалізації / А. М. Зубков, Ю. М. Косовцов, В. А. Юнда, А. А. Звонко, В. В. Атаманюк, Р. В. Бубенчиков (Україна); заявник Національна академія сухопутних військ імені гетьмана Петра Сагайдачного. – №а201800937 ; заявл. 01.02.18 ; опубл. 26.10.20, Бюл. № 20.
18. Комп'ютерне моделювання мультиспектральної цілефонової обстановки при самонаведенні ракети на наземну ціль / [А. М. Зубков, В. А. Юнда, Ю. М. Косовцов та ін.] // Наука і техніка Повітряних Сил Збройних Сил України. – 2016. – № 1 (22). – С. 151–155.

# МАТЕМАТИЧНЕ ТА КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ

## MATHEMATICAL AND COMPUTER MODELING

UDC 004.4:004.032.26

### COMPARISON OF SHORT-TERM FORECASTING METHODS OF ELECTRICITY CONSUMPTION IN MICROGRIDS

**Parfenenko Yu. V.** – PhD, Associate Professor of the Information Technology Department, Sumy State University, Sumy, Ukraine.

**Shendryk V. V.** – PhD, Associate Professor of the Information Technology Department, Sumy State University, Sumy, Ukraine.

**Kholiavka Ye. P.** – Postgraduate student, Information Technology Department, Sumy State University, Sumy, Ukraine.

**Pavlenko P. M.** – Doctor of science, Professor at Air Transportation Management Department, National Aviation University, Kyiv, Ukraine.

#### ABSTRACT

**Context.** The current stage of development of the electric power industry is characterized by an intensive process of microgrid development and management. The feasibility of using a microgrid is determined by the fact that it has a number of advantages compared to classical methods of energy generation, transmission, and distribution. It is much easier to ensure the reliability of electricity supply within the microgrid than in large energy systems. Energy consumers in a microgrid can affect the power balancing process by regulating their loads, generating, storing, and releasing electricity. One of the main tasks of the microgrid is to provide consumers with electrical energy in a balance between its generation and consumption. This is achieved thanks to the intelligent management of the microgrid operation, which uses energy consumption forecasting data. This allows to increase the efficiency of energy infrastructure management.

**Objective.** The purpose of this work is to develop short-term electricity consumption forecasting models for various types of microgrid electricity consumers, which will improve the efficiency of energy infrastructure management and reduce electricity consumption.

**Method.** The SARIMA autoregressive model and the LSTM machine learning model are used to obtain forecast values of electricity consumption. AIC and BIC information criteria are used to compare autoregressive models. The accuracy of forecasting models is evaluated using MAE, RMSE, MAPE errors.

**Results.** The experiments that forecast the amount of electricity consumption for the different types of consumers were conducted. Forecasting was carried out for both LSTM and AR models on formed data sets at intervals of 6 hours, 1 day, and 3 days. The forecasting results of the LSTM model met the forecasting requirements, providing better forecasting quality compared to AR models.

**Conclusions.** The conducted study of electricity consumption forecasting made it possible to find universal forecasting models that meet the requirements of forecasting quality. A comparative analysis of developed time series forecasting models was performed, as a result of which the advantages of ML models over AR models were revealed. The predictive quality of the LSTM model showed the accuracy of the MAPE of forecasting electricity consumption for a private house – 0.1%, a dairy plant – 3.74%, and a gas station – 3.67%. The obtained results will allow to increase the efficiency of microgrid management, the distribution of electricity between electricity consumers to reduce the amount of energy consumption and prevent peak loads on the power grid.

**KEYWORDS:** Microgrid, machine learning, LSTM model, AR model, forecasting, electricity consumption.

#### ABBREVIATIONS

AR is an autoregressive;

ARMA is an autoregressive moving average;

ARIMA is an autoregressive integrated moving average;

SARIMAX is a seasonal auto-regressive integrated moving average with exogenous factors;

ML is a machine learning;

LSTM is a long short-term memory;

MAE is a mean absolute error;

RMSE is a root-mean-square error;

MAPE is a mean absolute percentage error.

#### NOMENCLATURE

$y_t$  is the value of the time series at time  $t$ ;

$X_{k,t}$  is an exogenous variables;

$\beta_t$  is aregression parameters;

$\varphi_p$  is a non-seasonal autoregression;

$\Phi_p$  is a seasonal autoregression;  
 $\theta_q$  is a conditions of non-seasonal moving average;  
 $\Theta_q$  is an average seasonal conditions of movement;  
 $B^s$  is a lag operator;  
 $z_t$  is a noise;  
 $x_t$  is an entrance;  
 $f_t$  is a forgotten gate;  
 $i_t$  is an entrance gate;  
 $\tilde{C}$  is an updating the cell;  
 $C_t$  is a state of the cell;  
 $o_t$  is an output gate;  
 $h_t$  is an output;  
 $\tanh$  is an activation function;  
 $\hat{\sigma}^2$  is an expected variance;  
 $n$  is a number of input data;  
 $\hat{y}_t$  is a predictive value.

## INTRODUCTION

The current stage of the power industry development is characterized by the intensive process of microgrid development and its integration with the big centralized electricity grid. The term microgrid is used to refer to an integrated energy system of small power with distributed generators and energy consumers. Such an energy system, as a rule, is located on a small area [1].

The feasibility of using a microgrid is determined by the fact that it has several advantages compared to classical methods of energy generation, transmission, and distribution. In Microgrid, the energy produced is mainly used by local consumers, which ensures a reduction in losses associated with the transmission and distribution of energy by electric networks [2].

It is much easier to ensure the reliability of electricity supply within the microgrid than in large energy systems. Energy consumers in a microgrid can participate in the power balancing process by regulating their loads, generating, storing, and releasing electricity. One of the main tasks during microgrid management is to forecast electricity consumption to prevent peak loads on the power grid [3, 4].

The electricity consumption forecasting task is quite relevant due to the necessity to correctly load distribution in electrical networks, ensure their reliable operation and uninterrupted power supply to consumers [5].

Efficiency in forecasting electricity consumption, which is evaluated as the correspondence of the required and actual data, can be achieved by solving the task, formulated as follows: with the minimum number of resources, it is necessary to provide consumers with all the necessary electricity [6]. A similar task often arises for power grid organizations and large industrial enterprises. For enterprises, this problem is caused by the fact that organizations of this type must calculate the demand for electricity when it is generated or purchased on the

wholesale market [7, 8]. In addition, such an independent calculation can also be used as a factor in detecting commercial losses of electricity, because currently it is one of the rather serious problems at the stage of electricity transmission to end users.

Electricity consumers face the task of calculating the necessary load on the microgrid, which is influenced by various factors, such as climate, geographical location, time of day, socio-economic and other factors [16]. From the available data, it is necessary to select the most significant ones and make a forecast of energy consumption [17].

Energy sources in a microgrid system may be renewable sources of electricity generation from solar panels and wind turbines, as well as energy stored in large-capacity storage batteries. Microgrids can also combine multiple energy sources to provide consumers with uninterrupted access to electricity. The volume of electricity consumption is not a stable value, therefore, for the normal functioning of the microgrid, short-term forecasting is necessary to determine the future volumes of electricity consumption [7].

The electricity consumption forecast allows to reduce risks when making decisions about balancing the operation of the power system and reducing consumption by end users.

The total energy consumption of the region depends on internal changes for the enterprise, the sector of household consumers and the social sphere [8]. All the changes that occur in the demand for electric energy pose the task of maintaining the balance between production and consumption, because the energy service provider must fully satisfy the needs of electricity. To forecast energy consumption, a few stages must be completed [11]:

- carry out a graphic or descriptive analysis of the available input information about electricity consumption and factors affecting it;
- study the obtained time series;
- choose forecasting methods and make forecast models taking into account the influence of external factors;
- evaluate the received forecast values and choose the best forecasting model.

There is no standard approach to forecasting electricity consumption, as each consumer has its own specific characteristics [5]. Electricity consumption has cyclical, specific, and random components. Approximately 70–80% of all changes have cyclical trends. Also, one of the researched factors is the regularities of a functional nature [6]. These regularities include deviations explained by relatively well-known factors that are specific to each consumer. The third component of the forecast is random variation. When forecasting, these changes are probabilistic in nature [16].

Microgrid control and operation is carried out using specialized software. It is designed to monitor, control and optimize distributed energy for management and support of local IT infrastructure. For the uninterrupted operation of the Microgrid, the monitoring of current

electricity consumption and the forecasting for future time periods are required, which ensures the energy systems stability[8].

**The object of the study** is the process of improving the microgrid efficiency by forecasting electricity consumption for different types of electricity consumers.

**The subject of the study** are the time series forecasting methods that provide electricity consumption forecasting with high accuracy for various types of consumers.

**The purpose of the work** is the development of electricity consumption forecast models for various types of microgrid consumers', assessment of the forecast quality of forecasting models.

## 1 PROBLEM STATEMENT

To calculate the values of a time series at future points in time, it is necessary to define a functional dependence that reflects the relationship between the past and future values of this series [16]:

$$Z(\tau) = \Phi(Z(\tau-1), Z(\tau-2), Z(\tau-3), \dots) + \varepsilon_t. \quad (1)$$

Dependency (1) is called the forecasting model. It is necessary to create such a forecasting model for which the average absolute deviation of the true value from the forecast tends to the minimum for a given P [17]:

$$\bar{E} = \frac{1}{P} \sum_{t=T+1}^{T+P} |\varepsilon_t| \rightarrow \min. \quad (2)$$

The forecasted amount of electricity consumption depends on the values obtained in the previous time intervals of consumption.

Consumers of electricity in microgrid can be enterprises with a permanent or variable work schedule, household consumers, etc. This necessitates the task to create a universal forecasting model that will ensure high forecasting accuracy for all possible users of microgrid systems.

Short-term forecasting is used for decision support on the microgrid operation to choose the optimal mode of operation of the energy supply system, combining different types of renewable energy sources.

## 2 REVIEW OF THE LITERATURE

Many domestic and foreign scientists were engaged in solving the problem of electricity consumption forecasting as time series forecasting task including using machine learning methods. To generalize their experience, we studied some theoretical, methodological, and applied publications.

The high-precision data on electricity consumption by households were analysed in [9]. The authors separated the average demand profiles from demand fluctuations based solely on time series data and proposed a stochastic model for quantitative coverage of periodic demand fluctuations. The authors used the empirical mode

decomposition (EMD) to decompose the data into a finite number of functions based on the local properties of the data.

The authors of the [10] implemented machine learning (ML) models and described the application of machine learning for the development of energy collection (photovoltaics), storage (batteries), conversion (electrocatalysis), and control (smart grids).

In the study [11], a set of machine learning models of electricity consumption of a shoe store was built. As factors were taken: day of the week, day number, week number, holiday/working day, consumption of the previous day. The following machine learning methods were used in [11]: Linear Regression, Random Forest Regressor, Decision Tree Regressor, KNeighbors Regressor, LinearSVR. The model trained by the Random Forest Regressor method showed the best result.

The article [12] analyses the energy consumption of residential buildings. The model proposed by the authors generates a hidden space for demand peaks from the data fed into the long-short-term memory of a convolutional neural network (CNN-LSTM).

The authors of the article [13] used models based on the integration of seasonal autoregressive integrated moving average, Firefly optimization algorithm and support vector regression (SAMFOR). The comparison results showed that the SAMFOR model was more efficient than others such as Seasonal Autoregressive Integrated Moving Average (SARIMA) and Support Vector Regression (SVR) models, SARIMA-SVR and Random Forest (RF) models.

The authors of the article [20–21], on the basis of the future behavior of the microgrid system, predict the generation of electricity from the supply of energy sources. The authors model a system of microgrid with a large number of interchanges and changes for modeling the technology of generation of electricity and physical characteristics. In the article, there is a dynamic structure with mixed logic, to guarantee the future behavior for the microgrid, and also to protect the hour of the work of the accumulative battery. In order to predict inconsistencies in microgrid, a mechanism of reversal linkage is introduced to eliminate the problem of additional control of the horizon. With the help of model predictive control, the minimization of operating costs is achieved for planning the behavior of microgrid components.

The forecasting using artificial neural networks is widely used in modern research in the forecasting of electricity production and consumption [5]. Neural network models allow consideration of a large amount of data that must be processed in real-time and eliminate the uncertainty factor that further complicates the task of electricity consumption forecasting. Adaptation and constant learning makes the use of ML models the most effective for forecasting electricity consumption.

## 3 MATERIALS AND METHODS

The main factor that determines the power consumption regimes of the object supplied with



electricity in the microgrid is the nature of the electrical loads, which are generally divided into three types [3]:

- household – the load consumed by the population (residential buildings, dormitories);
- social – the load consumed by social objects (shops, schools, cinemas, etc.)
- industrial – the load consumed by enterprises with a permanent or variable work schedule.

To take into account all possible consumers, predictive models of electricity consumption are built on three data sets:

- household – a residential house;
- social – a gas station;
- industrial – a dairy plant.

Electricity consumption data was collected automatically every hour and entered into a database. Also, additional data such as ambient temperature, wind speed, and daylight length were used.

After a literature review, autoregressive models and machine learning models for electricity consumption forecasting were chosen. Nowadays they are the most popular time series forecasting models because, in most of the analysed articles, they showed better forecasting quality.

Autoregressive models are a statistical method of time series analysis. They use the linear dependence of the future value of the forecast on some number of previous values of the time series [3]. One of the main disadvantages of autoregressive models (AR, ARMA) is the inability to work with non-stationary data [10], therefore, the SARIMAX model was chosen. The SARIMAX model is a seasonal autoregressive integrated moving average model that considers exogenous variables that influence the determination of the forecast value.

$$y_t = \beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \dots + \beta_k X_{k,t} + \frac{(1-\theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)(1-\phi_1 B^S - \phi_2 B^{2S} - \dots - \phi_Q B^{QS})}{(1-\phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)(1-\Phi_1 B^S - \Phi_2 B^{2S} - \dots - \Phi_P B^{PS})} Z_t \quad (3)$$

A machine learning model such as long-short-term memory (LSTM) network was also chosen for short-term electricity consumption from microgrid (Figure 1). LSTM is a recurrent neural network (RNN) [15] for forecasting time series under conditions where peak values have uncertain periodicity. An advantage over other time series forecasting methods is the insensitivity to noise and outliers.

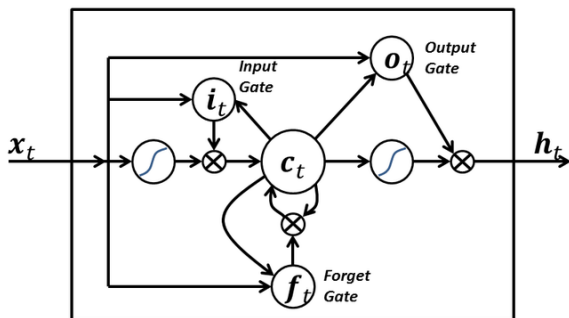


Figure 1 – LSTM Neural Network Architecture [22]

The principle of the model is as follows: first, the level of the forgetting filter determines what information can be forgotten or retained:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f). \quad (4)$$

Then the new information in the state of the cell is selected. In the next step, the input filter layer (sigmoid layer) updates the data. The tanh layer then creates a vector of new candidate values  $\tilde{C}_t$ , that can be added to the state of the cell:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \quad (5)$$

$$\tilde{C}_t = \text{than}(W_c \cdot [h_{t-1}, x_t] + b_c), \quad (6)$$

The next step is to replace the old state of the cell with a new one:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t, \quad (7)$$

The last step is to select the source information.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + o), \quad (8)$$

$$h_t = o_t \cdot \text{than}(C_t). \quad (9)$$

Information criteria are used to select the best parameters for the autoregression model. They allow you to compare models with each other and cannot test models in the sense of statistical hypotheses checking.

Akaïke (AIC) – the criterion for assessing the quality of statistical models, that is defined as follows:

$$AIC = \ln(\hat{\sigma}^2) + \frac{2(p+q+1)}{n}. \quad (10)$$

The Bayesian Information Criterion (BIC) is based on finding the maximum of the likelihood function. BIC is defined as follows:

$$BIC = \ln(\hat{\sigma}^2) + \frac{(p+q+1)\ln(n)}{n}. \quad (11)$$

To compare various forecasting models forecasting accuracy estimates are used.

Mean absolute error (MAE) is the arithmetic mean of absolute errors, calculated as follows:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|. \quad (12)$$

The root mean square error (RMSE) is the square root of the root mean square error, calculated as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|^2}. \quad (13)$$

The mean absolute percentage error (MAPE) is the mean absolute percentage deviation calculated as follows:

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \% \quad (14)$$

Forecasting errors are used to evaluate the quality of the developed forecasting models. Checking the accuracy of the built models allows to change the input parameters of the models to increase the forecasting accuracy. The forecasting model meets the accepted requirements if the value of the MAPE parameter is less than 5%.

#### 4 EXPERIMENTS

The study was carried out in several stages. First, several types of consumers with various sources of electricity consumption were selected, then different forecasting time intervals were considered. A microgrid system with renewable sources of electricity has been installed at all these consumers. Solar panels have been installed for a residential house and a gas station, and a wind generator for a dairy plant. Renewable sources partially cover the need for electricity.

A data set containing information on electricity consumption in a two-story private house was selected as a household consumer (Table 1).

Table 1 – An example of a data set for a residential house

Date	Value (kWh)	Day of week	notes	Hour	Month	Length of day
01.01.2018 0:00	1.057	2	weekday	0	June	49834
01.01.2018 1:00	1.171	2	weekday	1	June	49834
01.01.2018 2:00	0.560	2	weekday	2	June	49834
01.01.2018 3:00	0.828	2	weekday	3	June	49834
01.01.2018 4:00	0.932	2	weekday	4	June	49834

The level of electricity consumption in the house increases at night and during the hot season (Fig. 2).

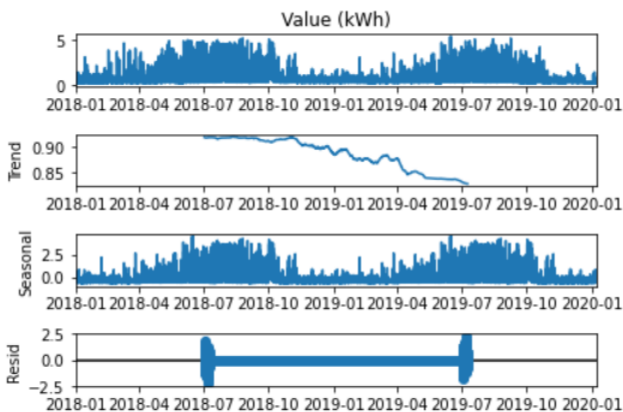


Figure 2 – Time series decomposition for residential consumer

The social electricity consumer is a gas station (Table 2).

Table 2 – An example of a data set for a gas station

Date	Value (kWh)	Indication	Length of day	Month	Week	Hour
01.01.2020 0:00	29.0	68.0	36190	11	44	0
01.01.2020 1:00	60.0	136.0	36190	11	44	1
01.01.2020 2:00	61.0	136.0	36190	11	44	2
01.01.2020 3:00	61.0	136.0	36190	11	44	3
01.01.2020 4:00	61.0	138.0	36190	11	44	4

The volume of electricity consumption at the gas station increases at night and in the cold season (Fig. 3).

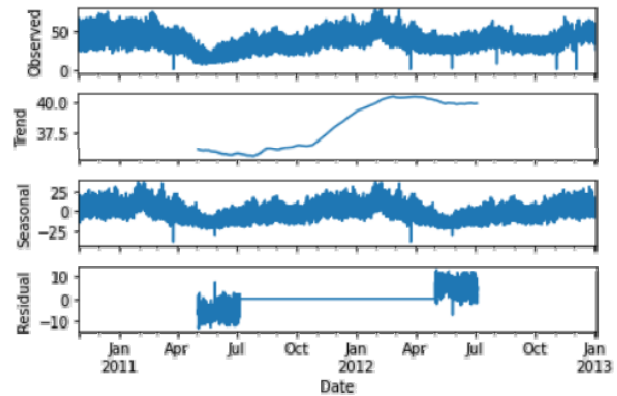


Figure 3 – Time series decomposition for gas station

Industrial electricity consumption data was provided by the dairy plant (Table 3).

Table 3 – An example of a data set for a dairy plant

Date	Value (kWh)	Length of day	T°	Week	Day of week	Hour
01.01.2013 0:00	223	42449	17	39	6	0
01.01.2013 1:00	215	42449	16	39	6	1
01.01.2013 2:00	218	42449	16	39	6	2
01.01.2013 3:00	210	42449	14	39	6	3
01.01.2013 4:00	214	42449	14	39	6	4

The amount of electricity used at the dairy plant increases during the day and during the hot season (Fig. 4).

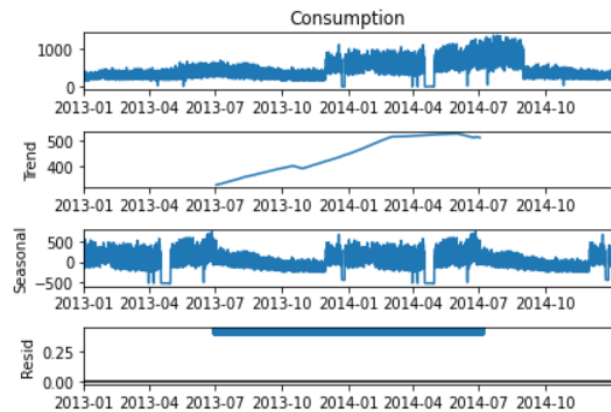


Figure 4 – Time series decomposition for dairy plant

Data from the above sources contained gaps that were filled with the average value of neighbouring indicators. The decomposition of the time series showed a seasonal

component, a large number of residuals, and the absence of a trend. Weather indicators of the environment, duration of a sunny day were taken as additional data.

The Python programming language and the Google Colab development environment were used to forecast the time series. The SARIMAX model was built for a residential house with the following parameters:  $p = 0$ ,  $d = 1$ ,  $q = 3$ ,  $P = 1$ ,  $D = 2$ ,  $Q = 1$ ,  $m = 12$ . The day of the week and the current time are chosen as exogenous parameters. The parameters which are shown in Table 4, provided the best values of the information criteria for a residential house.

Table 4 – Coefficients of the SARIMAX model for a residential house

	coef	std err	z	[0.025	0.975]
Day of week	0.0039	0.004	0.907	-0.004	0.012
Hour	0.0165	0.002	10.606	0.013	0.020
ar.L1	0.6891	0.004	162.102	0.681	0.697
ma.L1	-0.9856	0.001	-836.893	-0.988	-0.983
ar.S.L12	-0.2341	0.006	-40.045	-0.246	-0.223
ma.S.L12	-0.9450	0.002	-568.840	-0.948	-0.942
sigma2	0.2242	0.001	221.779	0.222	0.226

For the gas station, the SARIMAX model was built with the following parameters:  $p = 1$ ,  $d = 1$ ,  $q = 3$ ,  $P = 1$ ,  $D = 1$ ,  $Q = 1$ ,  $m = 12$ . The length of the day and the current time are additional data (Table 5).

Table 5 – Coefficients of the SARIMAX model for a gas station

	coef	std err	z	[0.025	0.975]
Length of day	0.0033	0.002	1.365	-0.001	0.008
Hour	0.0285	0.020	1.429	-0.011	0.068
ar.L1	0.7650	0.004	187.115	0.757	0.773
ma.L1	-1.0199	0.002	-616.789	-1.023	-1.017
ar.S.L12	-0.5938	0.004	-132.943	-0.603	-0.585
ma.S.L12	-1.2057	0.005	-251.137	-1.215	-1.196
sigma2	8.4633	0.083	102.044	8.301	8.626

For the dairy factory, the SARIMAX model was built with the following parameters:  $p = 1$ ,  $d = 0$ ,  $q = 1$ ,  $P = 1$ ,  $D = 1$ ,  $Q = 1$ ,  $m = 12$ . The length of the day and the ambient temperature serve as auxiliary data (Table 6).

Table 6 – Coefficients of the SARIMAX model for a dairy plant

	coef	std err	z	[0.025	0.975]
Length of day	0.0010	0.003	0.366	-0.004	0.006
T	2.2368	0.156	14.346	1.931	2.542
ar.L1	0.8849	0.004	246.114	0.878	0.892
ma.L1	-1.0000	0.022	-46.128	-1.042	-0.958
ar.S.L12	-0.4206	0.008	-55.627	-0.435	-0.406
ma.S.L12	-0.8688	0.004	-214.637	-0.877	-0.861
sigma2	860.8387	19.769	43.545	822.092	899.585

Another model for forecasting was the LSTM machine learning model. The model was built using the Keras library for Python. The model is built on four layers. First the Sequential class is instantiated, then the LSTM layers and the Dropout and Dense auxiliary layers are added. The number of neurons in the LSTM layer is indicated. Return\_sequences parameter is set to “true” to add the following data. The parameter Input\_shape indicates the number of time steps, and output\_shape shows the number of indicators. A dropout layer is added to avoid retraining. To ensure the reliability of the prediction, a Dense layer with the number of neurons 1 is added. The model for training uses 20 epochs, the data batch size is 100 (Table 7).

Table 7 – Coefficients of the LSTM model

Layer	Shape	Param
LSTM	100	52400
Dropout	100	0
Dense	1	101
Total params	52.501	
Trainable params	52.501	
Epochs	20	
Batch size	100	

Input values of the LSTM model for all types of consumers are indicators of hourly electricity consumption, the length of the day is used as an additional value. The coefficients of the LSTM model were selected in such a way that when the type of electricity consumers changed, the forecasted accuracy remained at a high level.

## 5 RESULTS

Experiments on forecasting volumes of electricity consumption for selected data sets were conducted. The predictive quality of the models was checked on test data. The selected models forecast hourly electricity consumption for 6 hours, 1 day, 3 days ahead.

The results of testing the forecasting accuracy of the SARIMAX model are in Table 8.

Table 8 – Forecasting accuracy of the SARIMAX model

Forecast period/ Assessment of accuracy		6 hours	1 day	3 days
Dairy Plant	MAE	36.6893	49.8139	64.2426
	RMSE	50.7658	110.2718	118.1284
	MAPE (%)	11.4544	13.1783	17.6610
Residential house	MAE	0.1145	0.2226	0.2393
	RMSE	0.1066	0.0043	0.3135
	MAPE (%)	0.0021	0.0070	0.0647
Gas Station	MAE	4.3309	4.6415	5.1155
	RMSE	11.5534	12.9961	5.9725
	MAPE (%)	5.6638	9.0617	11.5512

Forecasted by the SARIMAX model and the actual values of electricity consumption by the dairy plant are shown in Fig. 5.

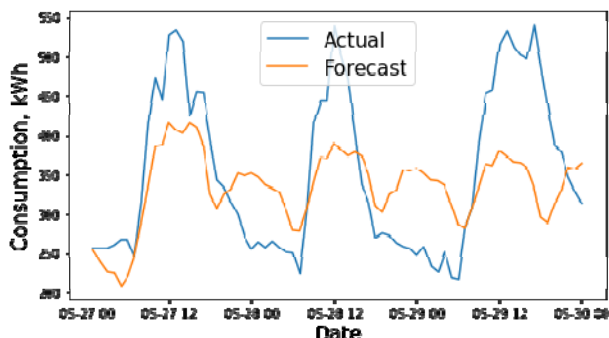


Figure 5 – Actual and forecasted values of electricity consumption by the dairy plant using the SARIMAX model

As a result of using the LSTM model built on parameters (Table 7), forecasting results (Table 9) were obtained for three types of electricity consumers.

Table 9 – Прогнозна якість моделі LSTM

Forecast period/ Assessment of accuracy		6 hours	1 day	3 days
Dairy Plant	MAE	15.4205	26.3461	24.5023
	RMSE	20.5534	35.6533	32.5615
	MAPE (%)	3.7420	6.7309	6.6436
Residential house	MAE	0.0902	0.0596	0.0908
	RMSE	0.1048	0.0784	0.2396
	MAPE (%)	0.1586	0.1631	0.1243
Gas Station	MAE	2.2020	3.0619	2.2143
	RMSE	2.4266	3.8967	2.9812
	MAPE (%)	3.6706	6.0989	4.6605

Forecasted by the LSTM model and the actual values of electricity consumption by the dairy plant are shown in Fig. 6.

The loss function (Fig. 7) was used to adjust the model, it is used to adjust the weights during the next evaluation of the model.

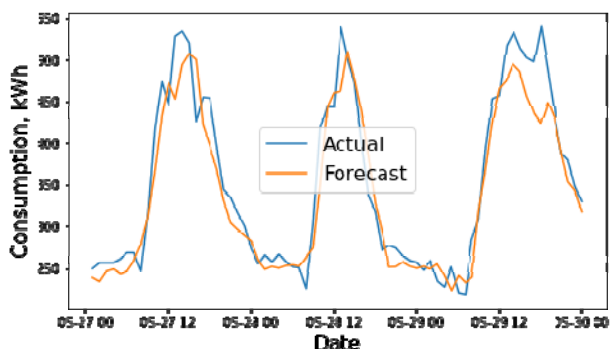


Figure 6 – Actual and forecasted values of electricity consumption by the dairy plant using the LSTM model

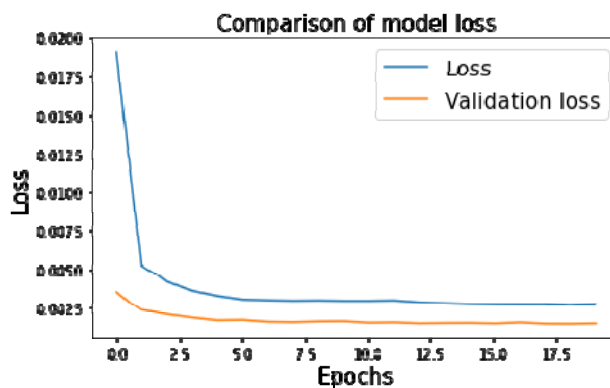


Figure 7 – Dependence of LSTM network loss on the number of training epochs on the dairy plant data set

## 6 DISCUSSION

The following conclusions can be drawn from the analysis of the results of the forecasting of electric energy in Microgrids. The SARIMAX model does not meet the requirement for forecasting electricity consumption. This model can only be used for short-term forecasting (from 1 hour to 6 hours). Only the prediction of electricity consumption for a private house using the SARIMAX model showed a slight advantage in predictive quality.

The forecasting results of the LSTM model met the forecasting requirements, providing better forecasting quality compared to AR models.

Compared to existing electricity forecasting works [7–12], forecasting results for the private sector have been significantly improved. The quality of forecasting of the LSTM model with the proposed parameters showed better results and forecasts than in the reviewed works. The proposed universal model showed high forecasting results for the considered types of consumers, compared to articles [10] where model parameters need to be constantly adjusted to obtain competitive forecasting results.

## CONCLUSIONS

A study of the forecasting of electricity consumption by various types of consumers was conducted, which made it possible to develop universal consumption forecasting models that meet the requirements of forecast quality.

Autoregressive models for selected data sets are built with different values of input parameters. Since the SARIMAX model is sensitive to the input data, there is a need for a large number of values (6–10 seasons) unlike the LSTM model. The parameters of the forecasting models were selected experimentally. The forecasting quality of each model was compared by conducting a series of experiments for each data set and for each time interval.

The comparative analysis of the developed time series forecasting models made it possible to draw a conclusion about the advantages of ML models over AR models. The forecasted quality of the LSTM model showed the accuracy of the MAPE of electricity consumption forecasting: for a residential house – 0.1%, a dairy – 3.74%, and a gas station – 3.67%.

**The scientific novelty** consists in building a universal forecasting model that provides high forecasting accuracy for different types of microgrid consumers.

**The practical significance** is that the developed forecasting models are used in decision support system for microgrid real-time operation. It allows effectively implement microgrid technology to increase the efficiency of energy infrastructure management and reduce electricity consumption.

**Prospects for further research** include the development of intelligent information technology to support decision-making and a set of software tools for energy infrastructure management, which will increase the efficiency of decision-making at various levels of energy infrastructure management.

#### ACKNOWLEDGEMENTS

The work was carried out with the support of the state budget research project of the Sumy State University “Intelligent Information-analytical Technologies and Means of Presentation, Assessment, and Management of the Country’s Energy Infrastructure” (state registration number 0121U109558).

#### REFERENCES

1. Issa Walid, Al-naemi Faris, Konstantopoulos George Stability analysis and control of a microgrid against circulating power between parallel inverters [Electronic resource], *Energy procedia*, 2019, Volume 157, pp. 1061–1070. Mode of access: <https://doi.org/10.1016/j.egypro.2018.11.273>. (date of access: 24.11.2022).
2. Sonam Parashar [et al.] Multiobjective optimal sizing of battery energy storage in grid-connected microgrid [Electronic resource], *The journal of engineering*, 2019, Vol. 2019, No. 18, pp. 5280–5283. Mode of access: <https://doi.org/10.1049/joe.2018.9237> (date of access: 24.11.2022).
3. Barzegarkhoo Reza, Farhangi Majid, Sze Lee Sing Nine-Level Nine-Switch Common-Ground Switched-Capacitor Inverter Suitable for High-Frequency AC-Microgrid Applications [Electronic resource], *IEEE transactions on power electronics*, 2022, Vol. 37, No. 5, pp. 6132–6143. Mode of access: <https://doi.org/10.1109/TPEL.2021.3131847> (date of access: 24.11.2022).
4. Shendryk Vira [et al.] Modeling techniques of electricity consumption forecasting [Electronic resource], *AIP conference proceedings*, 2022, No. 2570, 030004. Mode of access: <https://doi.org/10.1063/5.0100123> (date of access: 24.11.2022).
5. Dhahi Abu Electricity storage valuation framework. [Electronic resource]. IRENA, 2020. Mode of access: [https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2020/Mar/IRENA\\_storage\\_valuation\\_2020.pdf](https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2020/Mar/IRENA_storage_valuation_2020.pdf) (date of access: 24.11.2022).
6. Lianying Qi, Wanchun Dou, Wenping Wang Dynamic mobile crowdsourcing selection for electricity load forecasting [Electronic resource], *IEEE access*, 2018, pp. 256–263. Mode of access: <https://doi.org/10.1109/ACCESS.2018.2866641> (date of access: 24.11.2022).
7. Stanislaw Czapp, Hanan Tariq, Seweryn Szultka Electrical safety in low-voltage DC microgrids with B-type residual current devices [Electronic resource], *Eksplatacja i niezawodnosc – maintenance and reliability*, 2022, pp. 346–358. Mode of access: <https://doi.org/10.17531/ein.2022.2.16> (date of access: 24.11.2022).
8. Potapov Viktor, Khamitov Rustam, Makarov Vladimir Short-Term Forecast of Electricity Load for LLC “Omsk Energy Retail Company” Using Neural Network [Electronic resource], *Dynamics of Systems, Mechanisms and Machines*, 2018, pp. 1–5. Mode of access: <https://doi.org/10.1109/Dynamics.2018.8601430> (date of access: 24.11.2022).
9. Brancucci Carlo, Martinez-Anido Brancucci, Botor Benjamin, Florita Anthony The value of day-ahead solar power forecasting improvement [Electronic resource] // *Solar Energy*, 2016. pp. 192–203. Access mode: <https://doi.org/10.1016/j.solener.2016.01.049> (date of access: 24.11.2022).
10. Mehrmaz Anvari, Elisavet Proedrou, Benjamin Schäfer Data-driven load profiles and the dynamics of residential electricity consumption, *Nature Communications*, 2022, pp. 1711–1723. Access mode: <https://doi.org/10.1038/s41467-022-31942-9> (date of access: 24.11.2022).
11. Yao Zhenpeng, Lum Yanwei, Johnston Andrew Machine learning for a sustainable energy future [Electronic resource], *Nature reviews. Materials*, 2022, pp. 1–14. Access mode: <https://doi.org/10.1038/s41578-022-00490-5> (date of access: 24.11.2022).
12. Guillermo Hernández, Juan M. Corchado, Sigeru Omatu Machine Learning Models for Electricity Consumption Forecasting: A Review [Electronic resource], *International Conference on Computer Applications & Information Security*, 2019, pp. 1–6. Access mode: <https://doi.org/10.1109/CAIS.2019.8769508>. (date of access: 24.11.2022).
13. Hyung-Jun Moon, Seok-Jun Bu, Sung-Bae Cho Learning Disentangled Representation of Residential Power Demand Peak via Convolutional-Recurrent Triplet Network [Electronic resource], *International Conference on Data Mining Workshops (ICDMW)*, 2020, pp. 757–761. Access mode: <https://doi.org/10.1109/ICDMW51313.2020.00110> (date of access: 24.11.2022).
14. Ngoc-Tri Ngo, Anh-Duc Pham, Thi Thu Truong Developing a hybrid time-series artificial intelligence model to forecast energy use in buildings [Electronic resource], *Scientific Reports*, 2022, Vol. 12, No. 1, pp. 2045–2053. Access mode: <https://doi.org/10.1038/s41598-022-19935-6> (date of access: 24.11.2022).
15. Gers Felix, Cummins Fred Understanding LSTM Networks [Electronic resource]. <https://colah.github.io/>. Mode of access: <https://colah.github.io/posts/2015-08-Understanding-LSTMs/> (date of access: 24.11.2022).
16. Robert H. Shumway, David S. Stoffer Time series analysis and its applications (springer texts in statistics). [S. l.]. Springer, 2005, 562 p.
17. John Lu Z. Q. The elements of statistical learning: data mining, inference, and prediction [Electronic resource], *Journal of the royal statistical society: series A (statistics in society)*, 2010, Vol. 173, No. 3, pp. 693–694. Mode of access: [https://doi.org/10.1111/j.1467-985x.2010.00646\\_6.x](https://doi.org/10.1111/j.1467-985x.2010.00646_6.x) (date of access: 24.11.2022).
18. Asterios Papageorgiou [et al.] Climate change impact of integrating a solar microgrid system into the Swedish electricity grid [Electronic resource], *Applied energy*, 2020.

- Vol. 268, pp. 164–181. Mode of access: <https://doi.org/10.1016/j.apenergy.2020.114981> (date of access: 24.11.2022).
19. Pramode K. Verma, Mayssa El Rifai, Kam Wai Clifford Chan Application of the multi-stage protocol in IEEE 802.11i [Electronic resource], *Multi-photon quantum secure communication*. Singapore, 2018, pp. 143–159. Mode of access: [https://doi.org/10.1007/978-981-10-8618-2\\_9](https://doi.org/10.1007/978-981-10-8618-2_9) (date of access: 24.11.2022).
20. Sherin A. Helal [et al.] Energy management system for smart hybrid AC/DC microgrids in remote communities [Electronic resource], *Electric power components and systems*, 2019, Vol. 47, No. 11–12, pp. 1012–1024. Mode of access: <https://doi.org/10.1080/15325008.2019.1629512> (date of access: 24.11.2022).
21. Changbin Hu [et al.] Coordinative optimization control of microgrid based on model predictive control [Electronic resource], *Research anthology on smart grid and microgrid development*, [S. l.], 2022, pp. 233–252. Mode of access: <https://doi.org/10.4018/978-1-6684-3666-0.ch012> (date of access: 24.11.2022).
22. Kosarev V. S. Neural network approach to the electricity market forecasting with different-frequency data among the explanatory variables [Electronic resource], *2022 III international conference on neural networks and neurotechnologies (neuront)*. Saint Petersburg, Russian Federation, 16 June 2022, [S. l.], 2022. Mode of access: <https://doi.org/10.1109/neuront55429.2022.9805538> (date of access: 24.11.2022).
23. Kinanti Hanugera Gusti, Irhamah Irhamah, Heri Kuswanto Hybrid double seasonal ARIMA and support vector regression in short-term electricity load forecasting [Electronic resource], *IPTEK journal of proceedings series*, 2021, No. 6, pp. 322–341. Mode of access: <https://doi.org/10.12962/j23546026.y2020i6.11117> (date of access: 24.11.2022).

Received 10.11.2022.  
Accepted 13.12.2022.

УДК 004.4:004.032.26

## ПОРІВНЯННЯ МЕТОДІВ КОРОТКОСТРОКОВОГО ПРОГНОЗУВАННЯ СПОЖИВАННЯ ЕЛЕКТРОЕНЕРГІЇ ДЛЯ МІКРОМЕРЕЖ

**Парфененко Ю. В.** – канд. техн. наук, доцент кафедри інформаційних технологій, Сумський державний університет, Суми, Україна.

**Шендрюк В. В.** – канд. техн. наук, доцент кафедри інформаційних технологій, Сумський державний університет, Суми, Україна.

**Холявка Є. П.** – аспірант, кафедра інформаційних технологій, Сумський державний університет, Суми, Україна.

**Павленко П. М.** – д-р техн. наук, професор, професор кафедри організації авіаційних перевезень, Національний авіаційний університет, Київ, Україна.

### АНОТАЦІЯ

**Актуальність.** Сучасний етап розвитку електроенергетики характеризується інтенсивним процесом розвитку та управління мікромережею. Доцільність використання мікромережі визначається тим, що вона має низку переваг порівняно з класичними методами генерації, передачі та розподілу енергії. Забезпечити надійність електропостачання в мікромережі набагато легше, ніж у великих централізованих енергосистемах. Споживачі енергії в мікромережі можуть впливати на процес балансування електроенергії, регулюючи свої навантаження, генеруючи, накопичуючи та відпускаючи електроенергію. Одним із головних завдань Microgrid є забезпечення споживачів електричною енергією в балансі між її генерацією та споживанням. Це досягається завдяки інтелектуальному управлінню роботою Microgrid, яке використовує дані прогнозування енергоспоживання. Це дозволяє підвищити ефективність управління енергетичною інфраструктурою, робить її більш стійкою.

**Мета.** Метою даної роботи є розробка моделей короткострокового прогнозування споживання електроенергії для різних типів споживачів електроенергії у Microgrid, що дозволить підвищити ефективність управління енергетичною інфраструктурою та загалом зменшити споживання електроенергії.

**Метод.** Для отримання прогнозних значень споживання електроенергії використовуються авторегресійна модель (AR) SARIMA та модель машинного навчання (ML) LSTM. Інформаційні критерії AIC і BIC використовуються для порівняння авторегресійних моделей. Точність моделей прогнозування оцінюється за допомогою помилок MAE, RMSE, MAPE.

**Результати.** Проведено експерименти з прогнозування обсягів споживання електроенергії для різних типів споживачів. Прогнозування проводилося як з використанням моделей LSTM, так і моделей AR на сформованих наборах даних з інтервалами кожну годину протягом 6 годин, 1 день і 3 дні. Результати прогнозування з використанням моделі LSTM відповідали вимогам, забезпечуючи кращу якість прогнозування порівняно з авторегресійними моделями.

**Висновки.** Проведене дослідження прогнозування споживання електроенергії дозволило знайти універсальні моделі прогнозування, які відповідають вимогам якості прогнозування. Проведено порівняльний аналіз розроблених моделей прогнозування часових рядів, у результаті якого виявлено переваги моделей ML перед моделями AR. Прогностична якість моделі LSTM показала точність MAPE прогнозування споживання електроенергії для приватного будинку – 0,1%, молокозаводу – 3,74%, АЗС – 3,67%. Отримані результати дозволять підвищити ефективність управління мікромережею, розподілу електроенергії між споживачами для зменшення загальних обсягів споживання енергії та запобігання виникнення пікових навантажень.

**КЛЮЧОВІ СЛОВА:** мікромережа, машинне навчання, модель LSTM, модель AR, прогнозування, споживання електроенергії.

### ЛІТЕРАТУРА

1. Issa W. Stability analysis and control of a microgrid against circulating power between parallel inverters [Electronic resource] / Walid Issa, Faris Al-naemi, George Konstantopoulos // *Energy procedia*. – 2019. – Vol. 157. – P. 1061–1070. Mode of access: <https://doi.org/10.1016/j.egypro.2018.11.273>. (date of access: 24.11.2022).
2. Multiobjective optimal sizing of battery energy storage in grid-connected microgrid [Electronic resource] / Sonam Parashar [et al.] // *The journal of engineering*. – 2019. – Vol. 2019, No. 18. – P. 5280–5283. – Mode of access: <https://doi.org/10.1049/joe.2018.9237> (date of access: 24.11.2022).

© Parfenenko Yu. V., Shendryk V. V., Kholiavka Ye. P., Pavlenko P. M., 2023  
DOI 10.15588/1607-3274-2023-1-2

3. Barzegarkhoo R. Nine-Level Nine-Switch Common-Ground Switched-Capacitor Inverter Suitable for High-Frequency AC-Microgrid Applications [Electronic resource] / Reza Barzegarkhoo, Majid Farhangi, Lee Sze Sing // IEEE transactions on power electronics. – 2022. – Vol. 37, No. 5. – P. 6132–6143. – Mode of access: <https://doi.org/10.1109/TPEL.2021.3131847> (date of access: 24.11.2022).
4. Modeling techniques of electricity consumption forecasting [Electronic resource] / Vira Shendryk [et al.] // AIP conference proceedings. – 2022. – No. 2570. – 030004. – Mode of access: <https://doi.org/10.1063/5.0100123> (date of access: 24.11.2022).
5. Electricity storage valuation framework. [Electronic resource]. – Abu Dhabi : IRENA, 2020. – Mode of access: [https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2020/Mar/IRENA\\_storage\\_valuation\\_2020.pdf](https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2020/Mar/IRENA_storage_valuation_2020.pdf) (date of access: 24.11.2022).
6. Lianyong Q. Dynamic mobile crowdsourcing selection for electricity load forecasting [Electronic resource] / Qi Lianyong, Dou Wanchun, Wang Wenping // IEEE access. – 2018. – P. 256–263. – Mode of access: <https://doi.org/10.1109/ACCESS.2018.2866641> (date of access: 24.11.2022).
7. Stanislaw C. Electrical safety in low-voltage DC microgrids with B-type residual current devices [Electronic resource] / Czapp Stanislaw, Tariq Hanan, Szultka Seweryn // Eksploatacja i niezawodność – maintenance and reliability. – 2022. – P. 346–358. – Mode of access: <https://doi.org/10.17531/ein.2022.2.16> (date of access: 24.11.2022).
8. Potapov V. Short-Term Forecast of Electricity Load for LLC “Omsk Energy Retail Company” Using Neural Network [Electronic resource] / Viktor Potapov, Rustam Khamitov, Vladimir Makarov // Dynamics of Systems, Mechanisms and Machines. – 2018. – P. 1–5. – Mode of access: <https://doi.org/10.1109/Dynamics.2018.8601430> (date of access: 24.11.2022).
9. Brancucci C. The value of day-ahead solar power forecasting improvement [Electronic resource] / Carlo Brancucci Martinez-Anido, Benjamin Botor, Anthony Florita // Solar Energy. – 2016. – P. 192–203 – Access mode: <https://doi.org/10.1016/j.solener.2016.01.049> (date of access: 24.11.2022).
10. Anvari, M. Data-driven load profiles and the dynamics of residential electricity consumption / Anvari Mehrnaz, Proedrou Elisavet, Schäfer Benjamin // Nature Communications. – 2022. – P. 1711–1723. – Access mode: <https://doi.org/10.1038/s41467-022-31942-9> (date of access: 24.11.2022).
11. Yao Z. Machine learning for a sustainable energy future [Electronic resource] / Zhenpeng Yao, Yanwei Lum, Andrew Johnston // Nature reviews. Materials. – 2022. – P. 1–14. – Access mode: <https://doi.org/10.1038/s41578-022-00490-5> (date of access: 24.11.2022).
12. Hernández G. Machine Learning Models for Electricity Consumption Forecasting: A Review [Electronic resource] / Hernández Guillermo, Corchado Juan M., Omatu Sigeru // International Conference on Computer Applications & Information Security. – 2019. – P. 1–6. – Access mode: <https://doi.org/10.1109/CAIS.2019.8769508>. (date of access: 24.11.2022).
13. Moon H. Learning Disentangled Representation of Residential Power Demand Peak via Convolutional-Recurrent Triplet Network [Electronic resource] / Moon Hyung-Jun, Bu Seok-Jun, Cho Sung-Bae // International Conference on Data Mining Workshops (ICDMW). – 2020. – P. 757–761. – Access mode: <https://doi.org/10.1109/ICDMW51313.2020.00110> (date of access: 24.11.2022).
14. Ngo, NT. Developing a hybrid time-series artificial intelligence model to forecast energy use in buildings [Electronic resource] / Ngo Ngoc-Tri, Pham Anh-Duc, Truong Thi Thu // Scientific Reports. – 2022. – Vol. 12, No. 1. – P. 2045–2053. – Access mode: <https://doi.org/10.1038/s41598-022-19935-6> (date of access: 24.11.2022).
15. Gers F. Understanding LSTM Networks [Electronic resource] / Felix Gers, Fred Cummins // <https://colah.github.io/>. – Mode of access: <https://colah.github.io/posts/2015-08-Understanding-LSTMs/> (date of access: 24.11.2022).
16. Shumway R. H. Time series analysis and its applications (springer texts in statistics) / Robert H. Shumway, David S. Stoffer. – [S. l.] : Springer, 2005. – 562 p.
17. John Lu Z. Q. The elements of statistical learning: data mining, inference, and prediction [Electronic resource] / Z. Q. John Lu // Journal of the royal statistical society: series A (statistics in society). – 2010. – Vol. 173, No. 3. – P. 693–694. – Mode of access: [https://doi.org/10.1111/j.1467-985x.2010.00646\\_6.x](https://doi.org/10.1111/j.1467-985x.2010.00646_6.x) (date of access: 24.11.2022).
18. Climate change impact of integrating a solar microgrid system into the Swedish electricity grid [Electronic resource] / Asterios Papageorgiou [et al.] // Applied energy. – 2020. – Vol. 268. – P. 164–181. – Mode of access: <https://doi.org/10.1016/j.apenergy.2020.114981> (date of access: 24.11.2022).
19. Verma P. K. Application of the multi-stage protocol in IEEE 802.11i [Electronic resource] / Pramode K. Verma, Mayssaa El Rifai, Kam Wai Clifford Chan // Multi-photon quantum secure communication. – Singapore. – 2018. – P. 143–159. – Mode of access: [https://doi.org/10.1007/978-981-10-8618-2\\_9](https://doi.org/10.1007/978-981-10-8618-2_9) (date of access: 24.11.2022).
20. Energy management system for smart hybrid AC/DC microgrids in remote communities [Electronic resource] / Sherin A. Helal [et al.] // Electric power components and systems. – 2019. – Vol. 47, No. 11–12. – P. 1012–1024. – Mode of access: <https://doi.org/10.1080/15325008.2019.1629512> (date of access: 24.11.2022).
21. Coordinative optimization control of microgrid based on model predictive control [Electronic resource] / Changbin Hu [et al.] // Research anthology on smart grid and microgrid development. – [S. l.], 2022. – P. 233–252. – Mode of access: <https://doi.org/10.4018/978-1-6684-3666-0.ch012> (date of access: 24.11.2022).
22. Kosarev V. S. Neural network approach to the electricity market forecasting with different-frequency data among the explanatory variables [Electronic resource] / Vladimir S. Kosarev // 2022 III international conference on neural networks and neurotechnologies (neuront), Saint Petersburg, Russian Federation, 16 June 2022. – [S. l.], 2022. – Mode of access: <https://doi.org/10.1109/neuront55429.2022.9805538> (date of access: 24.11.2022).
23. Gusti K. H. Hybrid double seasonal ARIMA and support vector regression in short-term electricity load forecasting [Electronic resource] / Kinanti Hanugera Gusti, Irhamah Irhamah, Heri Kuswanto // IPTEK journal of proceedings series. – 2021. – No. 6. – P. 322. – Mode of access: <https://doi.org/10.12962/j23546026.y2020i6.11117> (date of access: 24.11.2022).

# НЕЙРОІНФОРМАТИКА ТА ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ

## NEUROINFORMATICS AND INTELLIGENT SYSTEMS

UDC 519.876.5:681.518.2

### COMPUTATIONAL INTELLIGENCE METHODS TO PATIENTS STRATIFICATION IN THE MEDICAL MONITORING SYSTEMS

**Bakumenko N. S.** – PhD, Associate Professor of Theoretical and Applied Systems Engineering Department, V. N. Karazin Kharkiv National University, Kharkiv, Ukraine.

**Strilets V. Y.** – PhD, Associate Professor of Theoretical and Applied Systems Engineering Department, V. N. Karazin Kharkiv National University, Kharkiv, Ukraine.

**Ugryumov M. L.** – Dr.Sc., Professor of Theoretical and Applied Systems Engineering Department, V. N. Karazin Kharkiv National University, Kharkiv, Ukraine.

**Zelenskiy R. O.** – Assistant of the Department of Urology and Pediatric Urology, Kharkiv Medical Academy of Postgraduate Education, Kharkiv, Ukraine.

**Ugryumova K. M.** – PhD, Associate Professor, Department of Mathematical Modeling and Artificial Intelligence, National Aerospace University “Kharkiv Aviation Institute”, Kharkiv, Ukraine.

**Starenkiy V. P.** – Dr.Sc., Professor, Head of the Radiation Oncology Department, State Organization “Grigoriev Institute for Medical Radiology and Oncology of the National Academy of Medical Sciences of Ukraine”, Kharkiv, Ukraine.

**Artiukh S. V.** – PhD, Senior Researcher of the Radiotherapy Group of the Radiology Department State Organization “Grigoriev Institute for Medical Radiology and Oncology of the National Academy of Medical Sciences of Ukraine”, Assistant Professor of Radiology and Radiation Medicine, Kharkiv National Medical University, Kharkiv, Ukraine.

**Nasonova A. M.** – PhD, Clinical Oncologist, Head of the Clinical Oncology and Hematology Department, State Organization “Grigoriev Institute for Medical Radiology and Oncology of the National Academy of Medical Sciences of Ukraine”, Kharkiv, Ukraine.

#### ABSTRACT

**Context.** In modern medical practice the automation and information technologies are increasingly being implemented for diagnosing diseases, monitoring the condition of patients, determining the treatment program, etc. Therefore, the development of new and improvement of existing methods of the patient stratification in the medical monitoring systems is timely and necessary.

**Objective.** The goal of intelligent diagnostics of patient’s state in the medical monitoring systems – reducing the likelihood of adverse states based on the choice of an individual treatment program:

- reducing the probability of incorrectly determining the state of the patients when monitoring patients;
- obtaining stable effective estimates of unknown values of treatment actions for patients (corresponding to the found state);
- the choice of a rational individual treatment program for the patients, identified on the basis of the forecasted state.

**Method.** Proposed methodology, which includes the following computational intelligence methods to patient’s stratification in the medical monitoring systems:

1) method of cluster analysis based on the agent-based approach – the determination of the possible number of patient’s states using controlled variables of state;

2) method of robust metamodels development by means artificial neuron networks under a priori data uncertainty (only accuracy of measurements is known) in the monitoring data: a) a multidimensional logistic regression model in the form of analytical dependences of the posterior probabilities of different states of the patients on the control and controlled variables of state; b) a multidimensional diagnostic model in the form of analytical dependences of the objective functions (quality criteria of the patient’s state) on the control and controlled variables of state;

3) method of estimating informativeness controlled variables of state at a priori data uncertainty;



4) method of robust multidimensional models development for the patient's state control under a priori data uncertainty in the monitoring data in the form of analytical dependencies predicted from the measured values of the control and controlled variables of state in the monitoring process;

5) method of reducing the controlled state variables space dimension based on the analysis of the variables informativeness of the robust multidimensional models for the patient's state control;

6) method of patient's states determination based on the classification problem solution with the values of the control and forecasted controlled variables of state with using the probabilistic neural networks;

7) method of synthesis the rational individual patient's treatment program in the medical monitoring system, for the state identified on the basis of the forecast.

Proposed the structure of the model for choosing the rational individual patient's treatment program based on IT Data Stream Mining, which implements the «Big Data for Better Outcomes» concept.

**Results.** The developed advanced computational intelligence methods for forecast states were used in choosing the tactics of treating patients, to forecast treatment complications and assess the patient's curability before and during special treatment.

**Conclusions.** Experience in the implementation of "Big Data for Better Outcomes" concept for the solution of the problem of computational models for new patient stratification strategies is presented. Advanced methodology, computational methods for a patient stratification in the medical monitoring systems and applied information technology realizing them have been developed. The developed methods for forecast states can be used in choosing the tactics of treating patients, to forecast treatment complications and assess the patient's curability before and during special treatment.

**KEYWORDS:** Information Technology Data Stream Mining, Medical Monitoring Systems, Machine Learning Methods, Mathematical Models and Methods for Patient Stratification.

### ABBREVIATIONS

ANN is an artificial neural network;

DM is a decision-makers;

CDSS is a computer decision support system;

IDS is an intellectual diagnostics of systems;

MFFN is a multilayer feedforward artificial neural network;

MMS is a medical monitoring system;

RBF is a radial-basis function neural network.

### NOMENCLATURE

$\vec{X}^{(0)}$  is an input vector of state parameters values;

$\vec{F}$  is a vector of state's target functions;

$P$  is a number of the elements in the dataset;

$H_0$  is a dimension of the input vector;

$H_{K+1}$  is a dimension of the output vector;

$\vec{F}^{(K+1)}$  is an output vector of the ANN model;

$K$  is a number of neuron in the hidden layer;

$X_j$  is a set of the elements from  $j^{\text{th}}$  cluster;

$\vec{x}_{jp}$  is  $p^{\text{th}}$  vector of the element in the  $j^{\text{th}}$  cluster;

$P_j$  is a number of elements in the  $j^{\text{th}}$  cluster;

$k^*$  is a number of clusters;

$d(\vec{x}_{jp}, \vec{c}_j)$  is an intra-cluster distance;

$d_1(\vec{x}_{jp}, \vec{c}_j)$  is Manhattan distance;

$MD^2$  is Mahalanobis distance;

$D_{KL}$  is Kullback-Leibler entropy;

$w_{jp}$  is a membership matrix;

$\vec{c}_j$  is a vector of cluster's centers;

$M(\vec{c}_j)$  is an average measure of the intra-cluster distance;

$LF(x)$  is a loss function;

$Z$  is a target set of pairs of cluster's number and number of elements in them;

$\hat{Z}$  is a solution of an optimize problem;

$k_j$  is the cluster's number;

$D_X$  is correctness set;

$\rho(\vec{x}_{jp}, \vec{c}_j)$  is Cauchy distribution;

$\eta$  is a parameter of Cauchy distribution;

$n$  is an iteration number;

$N$  is total number of iterations;

$k_t^{(n)}$  is estimated number of clusters on the  $n^{\text{th}}$  iteration;

$f_i$  is selection function's values at the ANN output;

$x_h$  is parameter's value at the input layer;

$MV$  is selection function of the MV-problem;

$\xi$  is significance level;

$f_{fit}$  is a fitness function;

$f_i^*$  is required value of the selection function;

$\sigma_{f_i, p}^0$  is standard deviation of the selection function;

$\beta_f$  is regularization parameter;

$n_\alpha$  is number of measurements for fixed state;

$\chi_{f_i, p}^2$  is chi-squared test;

$d_i$  is metric of the state's space;

$L_{fit}$  is convergence acceleration factor;

$\mu_i(f_i^*)$  is membership functions;

$f_i(\vec{X}^{(0)})$  is fitness function value;

$\sigma_{f_i}^0$  is standard deviation of variables value  $f_i$ ;

$f_{i, p}^*$  is required values of the function for  $p^{\text{th}}$  training set;

$f^0$  is normalized values of a variables or target functions;

$f$  is current values of a variables or target functions;

$f_{\min}$  is minimum value of a variables or target functions;

$f_{\max}$  is maximum value of a variables or target functions;

$l_f$  is slope factor of the activation function;

$bt$  is slope factor of the function  $th()$ ;

$\sigma_{X_h^{(0)}}^*$  is required standard deviation of the input data;  
 $X_{h,\max}^{(0)}$  is maximum value of the input data variables;  
 $X_{h,\min}^{(0)}$  is minimum value of the input data variables;  
 $D_L$  is training dataset;  
 $L$  is dimension of the training dataset (state's space);  
 $\hat{M}$  is search variables set;  
 $M$  is variables subset from  $D_L$  set;  
 $D_{F_i}$  is dispersion of the target function, in decibels;  
 $D_{F_i}^{(0)}$  is dispersion of the base (linear) model;  
 $D_{F_i}^{(model)}$  is dispersion of the other model;  
 $D_{M,\lambda}$  is subset of informativeness variables;  
 $\Sigma_X$  is the covariance matrix of variables  $X_l$  and  $X_n$ ;  
 $\sigma_{X_l}$  is the standard deviation of  $X_l$ ;  
 $r_{ln}$  are the correlation coefficients between  $X_l$  and  $X_n$   
( $l=1\dots L, n=1\dots L$ );  
 $E_i$  is a signal energy;  
 $\lambda_{ih}$  is the informativeness coefficients;  
 $D_Q$  is monitoring results;  
 $Q$  is dimension of monitoring data;  
 $\Pi^0$  is design and regime parameters;  
 $U^0$  is control variables;  
 $\Phi^0$  is phase variables;  
 $t$  is forecast moment;  
 $T_1$  is lower forecast moment;  
 $T_2$  is upper forecast moment;  
 $\varepsilon_l^0$  is relative error;  
 $\Psi$  is control state model;  
 $\Omega$  is diagnostic model;  
 $W^0$  is quality criteria of the patient's condition;  
 $I$  is dimension of time series  $q^0$ ;  
 $\vec{X}^*$  is vector of observed symptoms of unidentified precedent;  
 $X_1^*, \dots, X_A^*$  are principal components projections of  $\vec{X}^*$ ;  
 $A$  is number of the principal components;  
 $c_{lk}$  are the classes centers ( $l, k=1\dots A$ );  
 $p(R_{k*})$  is the probability distribution density agreeing to the Student's t-law;  
 $\langle \vec{X}_m \rangle$  is average projections value principal components vector of the observed symptoms of layer samples element;  
 $\Gamma(\cdot)$  is gamma function;  
 $K_l$  is number of precedents in  $l^{\text{th}}$  class,  $l=1\dots A$ ;  
 $K_k$  is number of precedents in  $k^{\text{th}}$  class,  $k=1\dots A$ ;  
 $\Sigma_{pooled}$  is the combined correlation matrix for the considered scenarios (for classes);  
 $P(R_k)$  is a prior probability of the classes realization;  
 $\hat{X}$  is quasisolution of MV-problem;  
 $t_f$  is Student's coefficient for considered function  $f$ ;

$Ro_f$  is Romanovsky's coefficient for considered function  $f$ ;  
 $M_\alpha[\cdot]$  is mathematical expectation with significance level  $\alpha$ ;  
 $x_p^*$  is required variables values  $x_p$  for prototype;  
 $\sigma_p^*$  is required mean deviation values of variables  $x_p$  for prototype;  
 $\sigma_{x_p}$  is mean square deviation of variables  $x_p \in X^0$ ;  
 $\sigma_{f_i}^*$  is the mean deviation values of decision making criterion  $f_i$  for prototype;  
 $\sigma_{f_i}$  is mean square deviation of decision making criterion  $f_i \in F$ ;  
 $\gamma$  is a regularization parameter;  
time is the value of the survival time for the patient after undergoing treatment.

## INTRODUCTION

The actual problems of modern medicine – the problems of assessing the states, forecasting the outcome of diseases, the effectiveness of treatment methods, assessing the likelihood of complications in patients – can be solved based on the use of advanced machine learning methods and information technologies that implement them. That's why developing an applied information technology for patient stratification in medical monitoring systems based on advanced machine learning methods is an actual scientific and practical task.

A medical monitoring system (MMS) consists of monitoring hardware for the patient's condition, decision makers (physicians) and computer decision support systems (CDSS). The aim of such systems – to ensure continuous observation, information collecting, data processing and analyzing patient's condition, forming recommendations for treating [1, 2].

National Aerospace University named by N. Zhukovsky "Kharkiv Aviation Institute" for more than 15 years has been developing applied information technology for patient stratification in medical monitoring systems based on machine learning methods. The university has developed its own software to accomplish the assigned problems and has qualified personnel who participate in international projects in the field of applied mathematics, statistics, and machine learning. In cooperation with the Kharkov National Medical University, the developed software is being tested for stratification of patients with prostate cancer, squamous cell carcinoma of the head and neck [3–14]. The developed methods for forecasted states are also planned to be verified on the data obtained in various oncological pathologies and to use them in choosing the tactics of treating patients, to forecast treatment complications and assess the patient's curability before and during special treatment.

The material and technical support and personnel base of the Kharkov National Medical University allows col-

lecting data on the course of treatment of cancer patients receiving chemoradiation treatment, using the whole range of non-invasive diagnostic methods, a wide range of tumor markers. Working with patients is regulated by the Bioethics and Deontology Committee and complies with international GCP standards.

As part of the implementation of projects, an innovative strategy is proposed for choosing a rational individual tactics for treating patients based on forecasted states identified using robust stratification methods, which will improve clinical results and prevent complications.

It is also planned to collect data on patients with cancer who have undergone COVID-19 before starting treatment and to assess the effect of the latter on the timing of the start of radiation and chemotherapy, depending on the severity of post Covid's syndrome.

This paper shows computational models for processing and research about a patient's condition and his body functioning based on the "Big data for better outcomes" concept for MMS.

**The object of study** is the processes for predicting the states and choosing individual treatment programs for patients based on monitoring data.

**The subject of study** is the computational intelligence methods to patient's stratification in the medical monitoring systems.

**The purpose of the work** is to reduce the likelihood of adverse states for patients in medical monitoring systems based on the choice of an individual treatment program by means of computational intelligence methods.

## 1 PROBLEM STATEMENT

Reducing material costs and time of forecasting of the patients' states in the medical monitoring systems are possible through the automation forecasting process of the patient's states using information technology Big Data. We define Big Data as an information technology based on the use of approaches series, tools and processing methods of the structured and unstructured data of large volumes and considerable variety to obtain the results for DM that are effective in conditions of continuous increase and distribution by numerous nodes of the computer network data stream.

The goals of intelligent diagnostics of patient's state in the medical monitoring systems:

- reducing the likelihood of adverse states based on the choice of an individual treatment program;
- reducing the probability of incorrectly determining the state of the patients (errors of the third kind in classifying the state of the systems) when monitoring patients;
- obtaining stable effective estimates of unknown values of treatment actions for patients (corresponding to the found state);
- the choice of a rational individual treatment program for the patients, identified on the basis of the forecasted state.

The aim of the investigation is to develop advanced methodology for solving the synthesis problem of patient's individual treatment program in the medical monitoring systems by means of computational intelligence methods for the medical information analysis and applied information technology realizing them by means of new robust estimation methodology (M-estimation) based on the concept of invariance of the theory of optimal control and apply it to solving nonlinear multidisciplinary problem of under uncertainty.

Let us input data  $(\bar{X}^{(0)}, \bar{F})_p, p = 1 \dots P$  are the results of the anamnesis (personal, visual examination, laboratory data), these are statistical data (dataset) from the observed patients, accumulated by medical institutions, as well as the individual results of patient monitoring, considered as training pairs. An input vector and output vector is of dimension  $H_0$  and  $H_{K+1}$  respectively.

The data model, such as information from the ambulatory card, formed by experts (physicians).

Statistical data (dataset) from the observed patients had transferred to the database storage Big Data from medical organizations.

Each patient must be able to contribute their own monitoring results to the database.

Data processing and diagnostic results transmission carried out by using Internet resources, which is available to every user, including the use of CDSS on remote servers.

Structuring new mathematical statements and developing computational methods is necessary for solving the synthesis problem of a patient's individual treatment program in the medical monitoring systems by means of computational intelligence methods, in stochastic formulation (MV-problem).

On the result of processing the input data, you need to find the following:

- the required number of controlled state variables for the classification of states (single or multiple diseases, and their corresponding disease stages);
- results of the patient's condition classification (single or multiple cases and their corresponding disease stages);
- estimates of survival rates and clinical effectiveness of the treatment;
- quality of life assessments for patients during treatment and in the follow-up period;
- results of informative monitored condition variables synthesis for patients individually at the current point in the monitoring process;
- results of predicting the time series of monitored state variables (including medical influences);
- justification of the time intervals choice for measuring values of controlled variables during monitoring depending on the disease stage;
- the results of predicting patient conditions that cannot be assessed using current clinical, laboratory and instrumental methods (e.g. hormone resistant prostate cancer);

– selection of an individual treatment regimen (medical interventions) depending on the patient's condition in order to minimize the risk of serious adverse events.

Such developments compared to existing methods will ensure a reduction in the probability of incorrectly determining the state of the systems (errors of the third kind in classifying the state of the patients), as well as obtaining stable effective estimations of the unknown values variables (corresponding to the found state).

## 2 REVIEW OF THE LITERATURE

Over the past decade, as a result of cooperation, the authors have developed the methods for solving such problems:

– formation of a subset of controlled variables of state, the values of the quantities of which are checked in by measuring instruments [3, 9];

– patient's classification of state results [4, 6, 7, 9];

– informativeness evaluations of the variables of for different stages of patients' diseases [4, 5, 9];

– robust metamodels: multidimensional diagnostic model, multidimensional logistic regression [6, 8, 9, 11];

– robust multidimensional models of control the states of the patients, evaluation of forecasted values of the controlled variables of state based on medical monitoring data for patients [11, 5];

– cluster analysis results – the number of recognizable states [5, 9, 12];

– patient's state classification results, using the values of the control and forecasted controlled variables of state [9, 11];

– the results of the synthesis of a rational individual treatment program for the patients for the state, determined on the basis of a forecast (determination of control variables (medical actions), that ensure the implementation of the treatment program) [9, 11, 14].

Numerical research was carried out with the help of the computer program "Non-linear evaluation methods in multicriteria problem of robust optimal designing and diagnostics of systems under parametric a priori uncertainty (methodology, methods, techniques and computer systems of support and decision-making implementing them)" (ROD&IDS) [13] developed by the authors.

The volumes of medical information about various diseases and their course are extremely large, and machine learning methods make it possible to process the accumulated information, take into account millions of different factors, social, territorial, demographic, genomic, etc., and make it possible to identify the unique features of each patient [15, 16]. It has been proven that the automation of the collection processes and further analysis of medical data allows to increase the accuracy of early diagnosis, prediction of the disease's development and the treatment effectiveness assessment [17]. For example, the Frost & Sullivan agency notes that artificial intelligence technologies increase the accuracy of diagnoses by 30–40%, and pathologist Andy Beck from Harvard Medical School believes that the further use of artificial

intelligence technologies will reduce the errors rate in diagnosis by 85% [18].

Modern technologies are also used to choose the most effective treatment strategy. Recently, scales (systems) for objectifying the assessment of clinical-physiological and laboratory parameters have been used to choose treatment tactics, the scope of anesthetic support and surgical intervention, predict the frequency of probable postoperative complications, lethality, and the treatment effectiveness [19]. The priority of such a strategy is that after receiving the sum of objective indicators, the doctor converts them into a score, which is ranked into numerical and staging corridors. This makes it possible to comprehensively assess the patient's condition at the moment, monitor his condition and carry out appropriate treatment. The correctness and effectiveness of treatment with this method depends on the qualifications and experience of the doctor treating the patient [20].

The choice of mathematical methods for the description and research of biological and medical objects depends both on the specialist's individual knowledge and on the specifics of the tasks to be solved [21].

The following methods are most often used to solve the classification problem of the medical facilities state:

– binary classification (decision trees and random forests) [22];

– artificial neural networks (ANN) [23];

– multidimensional logistic regression [24];

– naive Bayesian classifier [25];

– support vector machines [23, 26].

Obviously, the use of each of these methods separately does not allow solving the general problem of synthesizing an individual patient treatment program in a medical monitoring system.

So, there is a need to structure the system model of medical decision-making as a sequence of interrelated tasks and the corresponding system model of decision synthesis. In other words, it is necessary to create a methodology for the synthesis of solution to the problem of making medical decisions in general.

The work proposes methodology for solving the synthesis problem of a patient's individual treatment program in the medical monitoring systems by means of computational intelligence methods for the medical information analysis.

## 3 MATERIALS AND METHODS

Medical Monitoring System is a set of monitoring states hardware, information technology Big Data tools (which contains CDSS), patients and decision-makers (doctors) who are in a communicate relation with each other and united with the purpose of managing and organizing the process of systematic or continuous observation, collection, processing and information investigation about the object state (of patient), its functioning (of various organs) and development for a certain period of time. Its system is created and regulated by the monitoring entities (physicians) to ensure full, timely and accurate infor-

mation and appropriate organization of effective functioning and control of the object of diagnosis (patient).

The Context diagram of “Big Data for Better Outcomes” concept implementation in the medical monitoring systems is shown in Fig. 1.

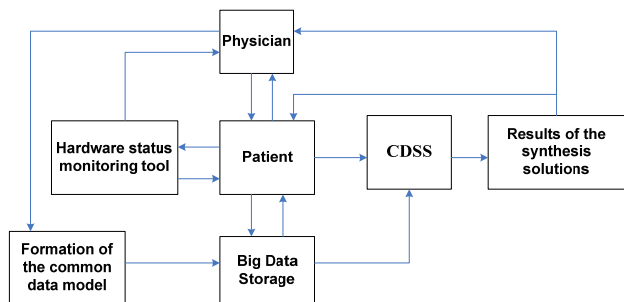


Figure 1 – Context diagram of “Big Data for Better Outcomes” concept implementation in the medical monitoring systems

Generalized methodology for solving the problem of synthesis problem of patient’s individual treatment program in the medical monitoring systems by means of computational intelligence methods as a result of its decomposition may be given as a sequence of processing techniques of structured and unstructured data of large volumes and considerable diversity using developed software. The solving result is sequence of steps to obtain mathematical model  $\vec{F}^{(K+1)}(\vec{X}^{(0)})$  in order to solve diagnosing problems.

Patient stratification is carried out in two steps: Data preprocessing and Monitoring the state of the patients. Each stage is described below.

1. Data Preprocessing (Preliminary data preparation is performed by a doctor):

1.1 Formation of a subset of controlled variables of state  $(\vec{X}^{(0)}, \vec{F})_p, p = 1...P$ , the values of the quantities of which are checked in by measuring instruments.

1.2 A total dataset of alternatives is generated. Each alternative includes subsets of control variables and controlled variables of state; objective functions. The total dataset includes subsets of alternatives corresponding to different states of the patients, including the healthy state.

1.3 Data cleaning from anomalous values of the quantities (outliers). Normalization of data.

1.4 Cluster analysis – the determination of the possible number of states (states of the patients) using control and controlled variables of state. We are looking for an answer to the question whether there is enough data from medical research to be able to recognize different states and the corresponding stages of diseases with the maximum a posteriori probability? If the number of identified clusters coincides with the specified one and the distances between the clusters are statistically significant, then the subset of monitored variables of state can be considered complete. Otherwise, the diagnostic system, which is used for monitoring of patient’s states, should be equipped with new measuring instruments.

Based on the agent-based approach and in accordance with the chosen measures of intra-cluster distance, permissive elite selection rules are proposed for the formation of clusters, the selection of the best of them, and also for the selection of elements into clusters in the process of solution synthesis.

The result of solving such a problem is the number of clusters, as well as the number of elements in them.

The c-means clustering method was chosen as the basis.

Let the sample of data be considered as  $X = \{X_j\}, X_j = \{\vec{x}_{jp}\}$  and  $P = \sum_{j=1}^{k^*} P_j^*$  is the total number of elements.

It is necessary to find  $\{k_j, P_j\}, j=1...k^*$ .

Four measures of intra-cluster distance were used as a metric for the clustering data:

$$d(\vec{x}_{jp}, \vec{c}_j) = \begin{cases} d_1(\vec{x}_{jp}, \vec{c}_j) \\ MD^2(\vec{x}_{jp}, \vec{c}_j) \\ w_{jp}^{-1} MD^2(\vec{x}_{jp}, \vec{c}_j) \\ -D_{KL}(\vec{x}_{jp}, \vec{c}_j) \end{cases}$$

Let us define the average measure of the intra-cluster distance:

$$M(\vec{c}_j) = \frac{1}{P_j} \sum_{p=1}^{P_j} d(\vec{x}_{jp}, \vec{c}_j).$$

Also the loss function is defined as

$$LF(X) = \frac{1}{k^*} \sum_{j=1}^{k^*} M(\vec{c}_j).$$

Then the **research problem statement** will take in the form:

$$\begin{cases} Z = (k_j, \{P_j\}) \\ \hat{Z} = \arg \min_{X \in D_x} LF(X). \end{cases}$$

It is necessary to determine the number of clusters and distribute the data among clusters so that the value of the loss function is minimal.

To correct the centers of clusters, we use the expression

$$\vec{c}_j = \frac{\sum_{p=1}^{P_j} w_{jp} \vec{x}_{jp}}{\sum_{p=1}^{P_j} w_{jp}},$$

where  $w_{jp} = \frac{\rho(\bar{x}_{jp}, \bar{c}_j)}{\sum_{p=1}^P \rho(\bar{x}_{jp}, \bar{c}_j)}$  is the membership matrix

with the Cauchy distribution

$$\rho(\bar{x}_{jp}, \bar{c}_j) = \frac{1}{\pi \eta \left[ 1 + \frac{MD^2(\bar{x}_{jp}, \bar{c}_j)}{\eta^2} \right]}$$

Data clustering algorithm based on agent-based approach:

1) set  $k_t^{(n)} > k^*$ ,  $P_j^{(n)} = \text{int}\left(\frac{N}{k_t^{(n)}}\right)$  and randomly

generate cluster centers  $\{\bar{c}_j\}$ ;

2) using the selected measure with cluster distance  $d_1(\bar{x}_{jp}, \bar{c}_j)$ , choose  $P_j^{(n)}$  the nearest neighbors for each  $j^{\text{th}}$  cluster;

3) for each  $j^{\text{th}}$  cluster, using the  $\{w_{jp}\}$  and  $\{\rho(\bar{x}_{jp} | P_j)\}$ , the cluster centers  $\{\bar{c}_j\}$  are corrected;

4) for each  $j^{\text{th}}$  cluster, using the selected measure within the cluster distance  $d(\bar{x}_{jp}, \bar{c}_j)$  and  $P_j^{(n)}$  the nearest neighbors are chosen. Delete the points that are duplicated:  $P_j^{(n)} \rightarrow P_j$ ;

5) for each  $j^{\text{th}}$  cluster the average measure of the intra-cluster distance  $M(\bar{c}_j)$  is calculated, and also  $LF(X)$  is calculated;

6) elite selection. Find the cluster with the largest  $M(\bar{c}_j)$  and delete it;

7)  $k_t^{(n+1)} = k_t^{(n)} - 1$ ;  $P_j^{(n+1)} = \text{int}\left(\frac{N}{k_t^{(n+1)}}\right)$ ;

8) back to step 2, if  $k_t^{(n+1)} > 1$ .

1.5 Development of robust metamodels at a priori data uncertainty in the monitoring data (given that the results of measurements of variables of states are random variables – only accuracy of measurements is known):

a) a multidimensional logistic regression in the form of analytical dependences of the posterior probabilities of different states of the patients on the control and controlled variables of state;

b) a multidimensional diagnostic model in the form of analytical dependences of the objective functions (quality criteria of the patient's state) on the control and controlled variables of state.

To development of robust metamodels as initial information used the vector function is given by a training sample  $(\bar{X}^{(0)}, \bar{F})_p, p=1\dots P$ . We must approximate the given set. The problem can be solved with a resultant

mathematical mechanism, which may give any value of the vector function  $\bar{F}_p^{(K+1)}(\bar{X}_p^{(0)})$ , represented by this training set at a fixed input vector within the range, limited by the input data.

A multilayer feedforward artificial neural network (MFFN) and radial basis function network (RBF), used for data approximation, is a parallel distributed processor, which is capable of saving acquired knowledge and processing information between local processor elements (neuro-elements or neurons), bound by special links (synaptic links).

To provide parameter stability (robustness) and informative capability of statistical systems and processes models on the basis of learning ANN at the a priori input data uncertainty and also practically sufficient data approximation, it is reasonable to use advanced deep learning methods – stable (robust) statistical assessment of their parameters with adaptive learning rate as the ANN learning method.

Student and V. I. Romanovsky are used as a smoothing functional when choosing a rational solution, which provides a stable (robust) estimation of the searching values with parametric uncertainty of the input data, as well as sufficient, from a practical point of view, accuracy of data approximation in problems of improving systems.

The function (MV-problem) was used as a scalar convolution of selection functions, considering

$f_i \equiv F_i^{(2)}, x_h = X_h^{(0)}$ :

$$MV = \frac{1}{2PI} \sum_{p=1}^P \xi^{P-p} \sum_{i=1}^I \left( f_{fit} \left[ \frac{\mu_i(f_{i,p}^*) \Delta_{f_{i,p}}}{f_i^* (1 + \sigma_{f_{i,p}}^0)} \right]^2 + \beta_f \cdot f_{fit} \left[ \frac{\chi_{f_{i,p}}^2 - n_\alpha}{\sqrt{2n_\alpha}} \right] \right)$$

here  $I=H_{k+1}$ ,  $\xi = [0.95, 0.99]$ ,

$$f_{fit}(d_i) = 1 - \exp\left[-\frac{L_{fit}}{4} d_i\right], L_{fit} \geq 4, (d_i > 0);$$

$$\Delta_{f_i} = F_i^{(K+1)}(\bar{X}^{(0)}) - f_i(\bar{X}^{(0)}), \sigma_{f_i}^0 = \frac{\sigma_{f_i}}{\sigma_{f_i}^*},$$

$$\left| \frac{\chi_{f_{i,p}}^2 - n_\alpha + 3}{\sqrt{2(n_\alpha - 3)}} \right| = \frac{n_\alpha}{\sqrt{2(n_\alpha - 3)}} \left| (\sigma_{f_i}^*)^2 - 1 + \frac{3}{n_\alpha} \right|, n_\alpha = 4.$$

For fitness functions  $f_i(\bar{X}^{(0)})$  in the expression for scalar convolution of selection functions  $MV$  the meanings of relative values are calculated by formulas:

– direct conversation  $f^0 = \frac{2l_f(f - \langle f \rangle)}{f_{\max} - f_{\min}}$ , here

$$\langle f \rangle = \frac{f_{\max} + f_{\min}}{2}, f^0 \in [-1, 1];$$

– inverse conversation  $f = \left[ (f_{\max} - f_{\min})f^0 / l_f + (f_{\max} + f_{\min}) \right] / 2$ , here  $l_f = th(bt)$  for MFFN and  $l_f = 1$  for RBF.

The relative mean square deviations of input data are calculated by formula

$$\left( \sigma_{f_i}^0 \right)^2 = \left( \frac{2l_f}{f_{i,\max} - f_{i,\min}} \right)^2 \left( \frac{\Delta_{f_i}^0}{300} f_{i,\max} \right)^2 n_\alpha,$$

$\Delta_{f_i}^0 = \frac{\Delta_{f_i}}{f_{i,\max}} 100\%$ , here  $l_f = th(\beta)$  for MFFN and  $l_f = 1$  for RBF;

$$\left( \sigma_{X_h^{(0)}}^* \right)^2 = \left( \frac{2}{X_{h,\max}^{(0)} - X_{h,\min}^{(0)}} \right)^2 \left( \frac{\Delta_{X_h^{(0)}}}{300} X_{h,\max}^{(0)} \right)^2 n_\alpha.$$

By known training set  $D_L$  the ANN parameters vector  $\hat{M}$  will choose according to the principle of maximum a posteriori probability distribution density:

$$\hat{M}_{n+1} = \arg \min_{M \in D_L} MV(M | D_L).$$

An advanced computational method is proposed for estimating the parameters of structural-parametric models of systems and processes in the form of trained ANNs based on the method of stochastic approximation and the use of a regularizing sequential (adaptive) algorithm for the synthesis of solutions with delayed correction (based on the ravine method of conjugate gradient methods and the method of simulating the movement of bee colonies), which implements adaptive control of calculations in accordance with the principle of minimum disturbance.

Application of the proposed methods avoids the appearance of false ravines or valleys on response surfaces in case of gross errors in the input data.

In the paired comparison of metamodels models changing of the signal variance is evaluating, which defines the robustness of a particular model:

$$D_{F_i}, dB = 10 \lg \left( \frac{D_{F_i}^{(model)}}{D_{F_i}^{(0)}} \right),$$

here  $model = 1$  for MFFN and  $model = 2$  for RBF.

1. Monitoring of the state of the patients (data processing and analysis is performed by the doctor together with the patient):

2.1 The values of the control and controlled variables of state corresponding to the patient's state at a given time are measured.

2.2 Determination of the state for which the maximum a posteriori probability of its realization corresponds with the observed values of the quantities based on the solution of the classification problem, which allows to determine the disease stages that are not recognized by modern biomarkers (e.g., hormone-resistant stage in prostate cancer).

2.3 Estimation informativeness controlled variables of state at a priori data uncertainty, the synthesis of the set of informative controlled variables of state according to the patient's state (disease stage) for the reduction of the variables of state space dimension, using multidimensional diagnostic model, i.e. searching informative subset  $D_{M,\lambda}$  of minimal dimension where  $D_{M,\lambda} \subset D_L$ .

The set of input data  $F_i(X)$ , where  $X = \{x_l\}$ ,  $l = 1 \dots L$  is presented as a Taylor series, while retaining only the terms of the first infinitesimal order in the expansion. For the dispersion of an arbitrary gotten linear function of several random variables estimate holds:

$$D_{F_i} = (grad F_i)^T \Sigma_X grad F_i = \dots = \sum_{l=1}^L \left( \frac{\partial F_i}{\partial x_l} \right)^2 \sigma_{x_l}^2 + \sum_{l=1}^L \sum_{n=1, n \neq l}^L r_{ln} \frac{\partial F_i}{\partial x_l} \frac{\partial F_i}{\partial x_n} \sigma_{x_l} \sigma_{x_n}.$$

Let us define the signal energy by the expression

$$E_i = \sum_{h=1}^{H_0} \left| D_{F_i^{(2)}} | X_h^{(0)} \right|,$$

$$D_{F_i^{(2)}} | X_h^{(0)} = \left( \frac{\partial F_i^{(2)}}{\partial X_h^{(0)}} \right)^2 \sigma_{X_h^{(0)}}^2 + \left( \sum_{n=1, n \neq h}^{H_0} r_{ln} \frac{\partial F_i^{(2)}}{\partial X_n^{(0)}} \sigma_{X_n^{(0)}} \right) \frac{\partial F_i^{(2)}}{\partial X_h^{(0)}} \sigma_{X_h^{(0)}}.$$

The informativeness coefficients (contribution weight  $X_h^{(0)}$  into  $F_i^{(2)}$ ) are defined by

$$\lambda_{ih} = \frac{\left| D_{F_i^{(2)}} | X_h^{(0)} \right|}{E_i}, \sum_{l=1}^L \lambda_{ih} = 1.$$

2.4 Development of robust multidimensional models of control the state for the patients at a priori data uncertainty in the monitoring data in the form of analytical dependencies predicted from the measured values of the control and controlled variables of state in the monitoring process.

2.5 Forecasting multidimensional time series of controlled variables of state based on multidimensional models of control of the state for the patients:

$$D_Q = \{q^0(t+1)\},$$

$$q^0(t+1) = [\Pi^0(t+1), U^0(t+1), \Phi^0(t+1)],$$

$$t_{lk} = \sqrt{\frac{MD_{lk}^2}{\frac{1}{K_l} + \frac{1}{K_k}}},$$

where  $t=T_1...T_2$  is a limited set of  $\Pi^0, U^0, \Phi^0$  and creating time series.

It is required to obtain a functional dependence, that will be reflect relationship between the next and previous values of the time series which satisfies the system preferences of DM, for a given forecast horizon:

$$q^0(t+1) = F(q^0(t+T_2-1), \dots, q^0(t-T_1)) + \varepsilon_t^0.$$

Controlled process mathematical model:

$$\Delta\Phi^0 = \Psi(\Delta\Phi^0, \Delta U^0)$$

$$\Delta W^0 = \Omega(\Delta\Phi^0, \Delta U^0)$$

where the first expression is recurrent mathematical model for monitoring and the second expression is a diagnostic model:

$$X^0(t+1) = \ln\left(\frac{q_i(t+1)}{q_i(t-T_1)}\right), i=1...I, t = (-T_1+1)...T_2.$$

2.6 Reducing the dimension of the space of controlled variables of state based on the analysis of the informativeness of the variables of the robust multidimensional models of control the state for the patients (Sensitivity Analysis). Estimating the rank of time series cointegration.

2.7 Determination of the states of the patients based on the solution of the classification problem and the values of the control and forecasted controlled variables of state is conducted using probabilistic neural networks. An input layer  $X_1^*, \dots, X_A^*$  elements are the principal component vector projections of the monitored state variables  $\vec{X}^*$ . The layer of samples  $c_{1k}, \dots, c_{Ak}$  are the classes centers of the training samples. The number of exemplars equals the number of classes in the training set. The layer of input and the samples are entirely meshed structure. An element activity of the samples layer was determined by the dependence that related to the probability distribution density agreeing to the Student's t-law (that is proper for the limited samples):

$$c_{lk} = \rho(R_k) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n} \Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{t_{lk}^2}{n}\right)^{-\frac{n+1}{2}},$$

here  $\Gamma$  is gamma function with  $n=K_l+K_k-2$ ;

here  $MD_{lk}^2$  is Mahalanobis distance from an unidentified precedent (it is supposed that it appropriate to  $l$  class) to the  $k$  sample:

$$MD_{lk}^2 = \frac{1}{A} (\vec{X}^* - \langle \vec{X}_k \rangle)^T \Sigma_{pooled}^{-1} (\vec{X}^* - \langle \vec{X}_k \rangle),$$

here  $\Sigma_{pooled}$  is determined because of belonging of the precedent to one, or another class:

$$\Sigma_{pooled} = \frac{1}{K_l + K_k - 2} ((K_l - 1)\Sigma_l + (K_k - 1)\Sigma_k).$$

The output layer  $k^*$ ,  $\rho(R_{k^*})$  is a discriminator of the threshold value indicating the class to which belongs an unknown precedent.

The values of t-Student statistics depend on the selection of the degree of proximity of the precedent to the samples (when solving the classification problem), as well as between the samples (when analyzing the significance of the distance between classes). Therefore, it is necessary to further forming of supplementary statistical decision rule selecting a single guide basis. According to the maximum likelihood rule as a criterion for the transition from one state to the other state of patients, Bayes formula is used:

$$\forall k = 0...A-1: \frac{\rho(\vec{X}_{k+1}^*)}{\rho(\vec{X}_k^*)} \geq 1.$$

Which is true provided that

$$\frac{P(R_k)}{P(R_{k+1})} \cdot \frac{\rho(\vec{X}_{k+1}^*)}{\rho(\vec{X}_k^*)} \approx 1.$$

2.8 Synthesis of a rational individual treatment program for the patients in the medical monitoring system, for the state identified on the basis of the forecast.

The control variables  $\hat{X} = (M[X_0], \sigma_x^0)$  estimation problem can be presented like a MV-problem.

The MV-problem provides an efficient stable (robust) estimation of the required quantities under parametric input data uncertainty.

The quasisolution of this problem is

$$\hat{X} = \arg \inf_{\hat{X} \in D_X} MV(\hat{X} | t_f, Ro_f),$$

where the scalar convolution of selection functions  $MV$  is



$$MV = \frac{1}{2I} \sum_{i=1}^I \left( f_{fit} \left[ \frac{\left( \frac{\mu_i(f_i^*) \Delta_{f_i}}{f_i^* (1 + \sigma_{f_i}^0)} \right)^2}{\left( \frac{\chi_{f_i}^2 - (n_\alpha - 3)}{\sqrt{2(n_\alpha - 3)}} \right)} \right] + \beta_f \cdot f_{fit} \left[ \frac{\left( \frac{\chi_{f_i}^2 - (n_\alpha - 3)}{\sqrt{2(n_\alpha - 3)}} \right)}{\left( \frac{\chi_{f_i}^2 - (n_\alpha - 3)}{\sqrt{2(n_\alpha - 3)}} \right)} \right] \right) + \gamma \frac{1}{2P} \sum_{p=1}^P \left( f_{fit} \left[ \frac{\left( \frac{\mu_p(x_p^*) \Delta_{x_p}}{x_p^* (1 + \sigma_{x_p}^0)} \right)^2}{\left( \frac{\chi_{x_p}^2 - (n_\alpha - 3)}{\sqrt{2(n_\alpha - 3)}} \right)} \right] + \beta_x \cdot f_{fit} \left[ \frac{\left( \frac{\chi_{x_p}^2 - (n_\alpha - 3)}{\sqrt{2(n_\alpha - 3)}} \right)}{\left( \frac{\chi_{x_p}^2 - (n_\alpha - 3)}{\sqrt{2(n_\alpha - 3)}} \right)} \right] \right).$$

Here  $\gamma=10^{-7} \dots 1.0$ ,  $I=H_{k+1}$ ,  $\Delta_{f_i} = M_\alpha[f_i] - f_i^*$ ,  
 $\Delta_{x_p} = M_\alpha[x_p] - x_p^*$ ,  $\chi_{f_i}^2 = n_\alpha \frac{M_\alpha[(f_i - M_\alpha[f_i])^2]}{(\sigma_{f_i}^*)^2}$ ,

$$\chi_{x_p}^2 = n_\alpha \frac{M_\alpha[(x_p - M_\alpha[x_p])^2]}{(\sigma_{x_p}^*)^2}, \sigma_{f_i}^0 = \frac{\sigma_{f_i}}{\sigma_{f_i}^*}, \sigma_{x_p}^* = \left\{ \begin{matrix} \sigma_{x_p} \\ \sigma_{x_p}^* \end{matrix} \right\},$$

$$\left| \frac{\chi_{f_i}^2 - (n_\alpha - 3)}{\sqrt{2(n_\alpha - 3)}} \right| = \frac{n_\alpha}{\sqrt{2(n_\alpha - 3)}} \left| (\sigma_{f_i}^*)^2 - 1 + \frac{3}{n_\alpha} \right|,$$

$$\left| \frac{\chi_{x_p}^2 - (n_\alpha - 3)}{\sqrt{2(n_\alpha - 3)}} \right| = \frac{n_\alpha}{\sqrt{2(n_\alpha - 3)}} \left| (\sigma_{x_p}^*)^2 - 1 + \frac{3}{n_\alpha} \right|.$$

For the MV-problem solution the computing method based on the memetic algorithm is applied. In this method the evolutionary algorithm was realized with the decremen-

tal neighborhood method and the randomize path laying method together. The evolutionary algorithm has next parameters, which change from epoch to epoch: real coding operator, fitness function and relaxation function.

The proposed method application gives the stable (robust) estimation of desired values under parametric input data uncertainty and lowers the difficulty of the quasisolution synthesis method.

The structure of the model for the choice of a rational individual treatment program for the patients based on IT Data Stream Mining, which implements the “Big Data for Better Outcomes” concept is shown in Fig. 2.

#### 4 EXPERIMENTS

In that work, as an implementation of the concept “Big Data for Better Outcomes” for stratification of patients, the following tasks were solved:

- formation of a subset of controlled variables of state, the values of the quantities of which are check in by measuring instruments;
- robust metamodels: multidimensional diagnostic model;
- informativeness evaluations of the variables of states for different stages of patients’ diseases.

The authors considered **the task of assessing the increased survival rate of cancer patients** who received chemoradiotherapy.

The multidimensional diagnostic model – model for estimating the survival of patients was built using a data set obtained in the real medical practice of the authors.

The patients with squamous cell carcinomas of the head and cancer (stages III, IVa, IVb) was included in study and did a comprehensive examination, including the collection of anamnestic data, general clinical physical examination, computed tomography of the head, neck and chest organs before the start of radiation therapy and 1 month after its completion, laboratory general clinical and biochemical blood tests. All patient received course of radiation therapy 5 Gy per fraction till 6–7 weeks (32–35

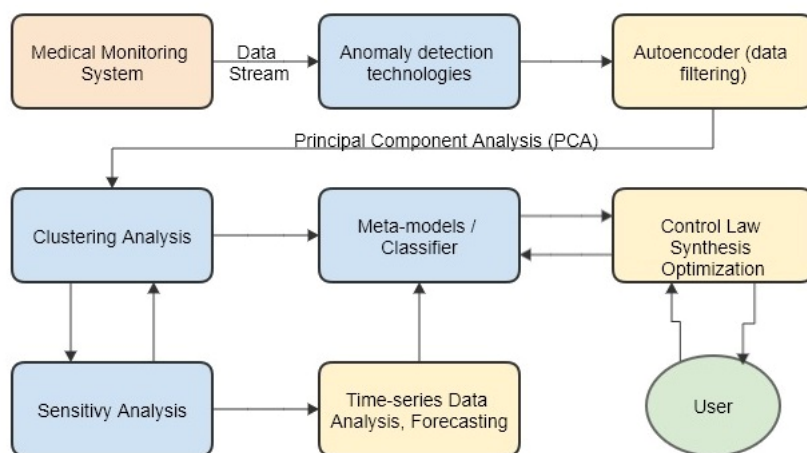


Figure 2 – Context diagram on the implementation of the model of the choice of a rational individual treatment program for the patients, identified on the basis of the forecasted state based on IT Data Stream Mining

fractions) with chemomodification Cisplatin 100 mg/m<sup>2</sup> at 1, 21, 42 days in SI “Grigoriev Institute for Medical Radiology and Oncology of NAMS of Ukraine”. This set contains records of 57 patients on 9 indicators (parameters) that characterize the current state of each patient: age, localization, stage, tumor (T) and lymph (N) nodes status (TNM classification, 8 edition), response (per RECIST 1.1 criteria), tumor marker squamous cell carcinoma’s antigen “before” and “after” the course treatment, survival time. All indicators have a numerical type, but some of them are enumerable, for example, the “localization” indicator takes four values from 1 to 4. The objective function for each record is “survival time”, the other parameters are attributed to the input data.

First, it is necessary to normalize the data. Since the inputs are different in value and content, the normalization for them will be different. The parameters “localization”, “stage”, “T” and “response” are normalized according to the formula

$$f_i^0 = \frac{f_i}{f_{\max}},$$

here  $f_{\max} = 4$ .

The age indicator is normalized as follows:

$$Age_{norm} = \frac{Age_{\max} - Age_{current}}{Age_{\max}},$$

here we take age as the maximum value  $Age_{\max} = 85$ .

The indicators of tumor markers “before” and “after” are reduced to one parameter of tumor markers as follows:

$$Oncomarker = \ln \frac{marker\_after}{marker\_before}.$$

Let’s denote the objective function by  $Suv\_time$ , its values are also pre-normalized by the formula

$$Suv\_time = \frac{time}{target\_time},$$

here  $target\_time = 25$ .

The results of the patient survival assessment are a functional relationship between the input data and the survival time. Using a robust meta-model based on a radial-basis neural network, survival time estimation values are calculated that correspond to the input data.

The structure of the simplest radial-basis function neural network includes three types of neuron layers. RBF with one hidden layer ( $K=1$ ). At the input layer RBF has  $H_0=7$  neurons, at the hidden layer –  $H_1 = 44$  neurons and  $H_2=1$  network outputs.

In addition, the task of determining the most influential (informative) parameters for assessing patient survival was solved.

Numerical research was carried out with the help of the computer program “Non-linear evaluation methods in multicriteria problem of robust optimal designing and intelligence diagnostics of systems under parametric a priori uncertainty (methodology, methods, techniques and computer systems of support and decision-making implementing them)” (ROD&IDS) developed by the authors [13].

## 5 RESULTS

To train the neural network and test its operation, the data set of the patient’s survival time was divided into two parts: training (47 records, 82%) and test (10 records, 18%).

As a result of training, the relative error was 1.5%.

The relative error on the test sample was 15.4%. The trained neural network can then be used to provide survival estimates for new patients (Table 1).

It was found by means of sensitive analysis of variables of the model for each patient from the training sample that on the average the indicators: T, age of the patient, response and stage of the disease are more likely to be informative (Table 2).

Table 2 – The informativeness evaluations of the indicators

	Age	Local.	Stage	T	N	Response	Tumor marker
Suv. time	0.262	0.0025	0.251	0.185	0.0006	0.289	0.009

Therefore, to assess the survival of patients, it is necessary to pay attention to these indicators.

Table 1 – Result of model for test set

Age	Local.	Stage	T	N	Response	Tumor marker	Suv_time (target)	Suv_time (fact)
0.235294	0.75	0.5	1	0.75	0.25	-0.887762	1	1.112889
0.164705	1	1	0	1	0.75	-0.062609	0.28	0.261637
0.364705	1	0.75	2	1	0.5	-0.243848	0.52	0.682176
0.258823	0.25	0.5	1	0.75	0.75	-0.267833	0.44	0.385309
0.352941	0.5	1	0	1	0.75	0.2626409	0.32	0.259509
0.341176	0.75	1	1	1	0.5	-1.659763	0.44	0.387156
0.305882	0.5	0.25	2	1	0.5	-0.212148	0.6	0.591057
0.188235	0.25	0.5	1	0.75	0.25	-0.472130	1	1.026191
0.258823	0.75	0.75	2	1	0.75	-0.819898	0.36	0.512091

## 6 DISCUSSION

Thus, the goal of the study was achieved – developed advanced methodology for solving the synthesis problem of the patient's individual treatment program in the medical monitoring systems by means of computational intelligence methods for the medical information analysis and applied information technology realizing them have been developed.

Its application allows:

- reducing the likelihood of adverse states based on the choice of an individual treatment program;
- reducing the probability of incorrectly determining the state of the patients (errors of the third kind in classifying the state of the systems) when monitoring patients;
- obtaining stable effective estimates of unknown values of treatment actions for patients (corresponding to the found state);
- the choice of a rational individual treatment program for the patients, identified on the basis of the forecasted state.

The developed software, which is described below information technology, is being tested for stratification of patients with prostate cancer, squamous cell carcinoma of the head and neck [3–11]. The developed methods for forecasted states are also planned to be verified on the data obtained in various oncological pathologies and to use them in choosing the tactics of treating patients, to forecast treatment complications and assess the patient's curability before and during special treatment.

The material and technical support and personnel base of the Kharkov National Medical University allows collecting data on the course of treatment of cancer patients receiving chemoradiation treatment, using the whole range of non-invasive diagnostic methods, a wide range of tumor markers. Working with patients is regulated by the Bioethics and Deontology Committee and complies with international GCP standards.

As part of the implementation of projects, an innovative strategy is proposed for choosing a rational individual tactics for treating patients based on forecasted states identified using robust stratification methods, which will improve clinical results and prevent complications.

It is also planned to collect data on patients with cancer who have undergone COVID-19 before starting treatment and to assess the effect of the latter on the timing of the start of radiation and chemotherapy, depending on the severity of post Covid's syndrome.

## CONCLUSIONS

**The scientific novelty** obtained results in the fact that an advanced methodology for solving the synthesis problem of a patient's individual treatment program in the medical monitoring systems by means of computational intelligence methods for the medical information analysis and applied information technology realizing them have been developed.

**The practical significance** of obtained results is that the advanced computational intelligence methods for forecast states can be used in choosing the tactics of treat-

ing patients, to forecast treatment complications and assess the patient's curability before and during special treatment.

**Prospects for further research** are to study application possibilities of the proposed methodology for the choice of a rational individual treatment program for the patients with various chronic diseases, identified on the basis of the forecasted state by means of IT Data Stream Mining.

## ACKNOWLEDGEMENTS

The work is supported by the state budget scientific research project of Karazin Kharkiv National University "Information process modeling in the complex and distributed systems" (state registration number 0121U109183) and research projects of Grigoriev Institute for Medical Radiology and Oncology of the National Academy of Medical Sciences of Ukraine "Development of programs for personalized control of the absorbed dose during radiation therapy of genital, head and neck tumors using in vivo dosimetry" (state registration number 0117U001046, 2017 – 2019) and "Optimizing topometrical preparation for radiation therapy of head and neck cancer patients" (state registration number 0119U103013, 2020 – 2022).

## REFERENCES

1. Conejar R. J., Kim H.-K. A Medical Decision Support System (DSS) for Ubiquitous Healthcare Diagnosis System, *International Journal of Software Engineering and Its Applications*, 2014, Vol. 8, № 10, pp. 237–244. DOI: <http://dx.doi.org/10.14257/ijseia.2014.8.10.22>.
2. Wagholikar K. B., Sundararajan V., Deshpande A. W. Modeling Paradigms for Medical Diagnostic Decision Support: A Survey and Future Directions, *Journal of Medical Systems*, 2012, Vol. 36, Issue 5, pp. 3029–3049. DOI: <https://doi.org/10.1007/s10916-011-9780-4>.
3. Starenkiy V., Goryachaya V., Sokolov O., Ugryumova E. Diagnostic model and information technology of classification states in the differential diagnosis NSCLC (non small cell lung cancer) patients with different methods of radiotherapy and chemotherapy, *Journal of Health Sciences*, 2013, Vol. 3 (8), pp. 7–26. ISSN: 1429-9623.
4. Goriacha V. A., Sokolov O. Yu., Ugryumova K. M., Antonyan I. M., Roshin Y. V., Zelensky A. I., Moshel F. G., Nalbandian T. A. Forecasting of Patients Condition in the \Monitoring Medical Systems on the Example of Prostate Diseases. *Journal of Education, Health and Sport*, 2016, Vol. 6, No. 5, pp. 77–93. DOI: <http://dx.doi.org/10.5281/zenodo.51160>.
5. Bakumenko N., Strilets V., Ugryumov M. Application of the C-Means Fuzzy Clustering Method for the Patient's State Recognition Problems in the Medicine Monitoring Systems [Electronic resource], *3rd International Conference on Computational Linguistics and Intelligent Systems (COLINS-2019) proceedings*, 2019, Vol. I: Main Conference, 10 p. Access mode: <http://ceur-ws.org/Vol-2362/paper20.pdf>
6. Strilets V., Bakumenko N., Chernysh S., Ugryumov M., Donets V. Application of artificial neural networks in the problems of the patient's condition diagnosis in medical monitoring systems, *Integrated Computer Technologies in Mechanical Engineering – Synergetic Engineering. Interna-*

- tional Scientific and Technical Conference (ICTM 2019)*, 2020, pp. 173–185. DOI: [https://doi.org/10.1007/978-3-030-37618-5\\_16](https://doi.org/10.1007/978-3-030-37618-5_16).
7. Strilets V., Bakumenko N., Donets V., Chernysh S., Ugryumov M., Goncharova T. Machine Learning Methods in Medicine Diagnostics Problem, *16th International Conference on ICT in Education, Research and Industrial Applications. Integration, Harmonization and Knowledge Transfer: proceedings*, 2020, Vol. II: Workshops, pp. 89–101. Access mode: <https://ceur-ws.org/Vol-2732/20200089.pdf>.
  8. Bakumenko N., Strilets V., Menailov I., Chernysh S., Ugryumov M., Goncharova T. Synthesis Method of Robust Neural Network Models of Systems and Processes, *Integrated Computer Technologies in Mechanical Engineering*. Springer, Cham, 2020, pp. 3–16. DOI: [https://doi.org/10.1007/978-3-030-66717-7\\_1](https://doi.org/10.1007/978-3-030-66717-7_1).
  9. Starenkiy V. P., Artiukh S. V., Zelenskiy O. I. and others *Patient's stratification in the medical monitoring systems*: monograph. Kharkiv National Medicine University, 2021, 148 p. (in Ukrainian)
  10. Mitryaeva N. A., Starenkiy V. P., Billozor N. V., Grebinyk L. V., Artyukh S. V. Radiosensitization by cyclooxygenase-2 inhibitors in radiation therapy of malignant neoplasms: monograph, *State Institution "Institute of Medical Radiology and Oncology named SP Grigorieva National Academy of Medical Sciences of Ukraine"*, 2021, 136 p.
  11. Starenkiy V., Artiukh S., Ugryumov M., Strilets V., Chernysh S., Chumachenko D. A Method for Assessing the Risks of Complications in Chemoradiation Treatment of Squamous Cell Carcinoma of the Head and Neck, *The Open Bioinformatics Journal*, 2021, Vol. 14, pp. 139–143. DOI: [10.2174/18750362021140100138](https://doi.org/10.2174/18750362021140100138)
  12. Strilets V., Donets V., Ugryumov M., Zelenskiy R., Goncharova T. Agent-Oriented data clustering for medical monitoring, *Radioelectronic and Computer Systemsthis*, 2022, № 1, P. 103–114. DOI: <https://doi.org/10.32620/reks.2022.1.08>
  13. Ugryumov M. L., Meniaylov Y. S., Chernysh S. V., Ugryumova K. M. (Ukraine) Computer program «Nonlinear estimation methods in the multicriterion problems of system's robust optimal designing and diagnosing under parametric apriority uncertainty (methodology, methods and computer decision support and making system)» («ROD&IDS»): Copyright registration certificate № 82875. Copyright and related rights. Official bulletin. Ministry of Economic Development and Trade of Ukraine. 2018, № 51, P. 403.
  14. Strilets V. E., Shmatkov S. I., Ugryumov M. L. et al. *Methods of machine learning in the problems of system analysis and decision making*: monograph. Karazin Kharkiv National University, 2020, 195 p. ISBN 978-966-285-627-9. (in Ukrainian)
  15. Sidey-Gibbons J. A. M., Sidey-Gibbons Chris J. Machine Learning in medicine: practical introduction, *BMC Medical Research Methodology*, 2019, No. 19. DOI: [10.1186/s12874-019-0681-4](https://doi.org/10.1186/s12874-019-0681-4).
  16. Varoquaux G., Cheplygina V. Machine Learning for medical imaging: methodological failures and recommendations for the future, *Digit. Med.*, 2022, No. 5 (48). DOI: <https://doi.org/10.1038/s41746-022-00592-y>.
  17. Antoniou T., Mamdani M. Evaluation of machine learning solutions in medicine, *CMAJ: Canadian medical association Journal*, 2021, No. 193 (36), E1425–E1429. DOI: [10.1503/cmaj.210036](https://doi.org/10.1503/cmaj.210036).
  18. Seyhan A. A., Carini C. Are innovation and new technologies in precision medicine paving a new era in patients centric care? *Journal of Translational Medicine*, 2019, No. 17, article number 114. DOI: <https://doi.org/10.1186/s12967-019-1864-9>
  19. Nayyar Anand, Gadhavi Lata, Zaman Noor. Chapter 2 – Machine learning in healthcare: review, opportunities and challenges, *Machine Learning and the Internet Medical Things in Healthcare*, 2021, P. 23–45. DOI: <https://doi.org/10.1016/B978-0-12-821229-5.00011-2>.
  20. Trujillano J., Badia M., Serviá L. et al. Stratification of the severity of critically ill patients with classification trees, *BMC Medical Research Methodology*, 2009, Vol. 9, No. 83. DOI: <https://doi.org/10.1186/1471-2288-9-83>.
  21. Amigo J. M., Small M. Mathematical methods in medicine: Neuroscience, cardiology and pathology, *Philosophical Transactions of The Royal Society A Mathematical Physical and Engineering Sciences*, 2017, Vol. 375 (2096): 20170016. DOI: [10.1098/rsta.2017.0016](https://doi.org/10.1098/rsta.2017.0016).
  22. Azar A., El-Metwally S. Decision tree classifiers for automated medical diagnosis, *Neural Computing and Applications*, 2013, Vol. 23, pp. 2387–2403. DOI: [10.1007/s00521-012-1196-6](https://doi.org/10.1007/s00521-012-1196-6).
  23. Bahari N. I. S., Ahmad A., Aboobaid B. M. Application of support vector machine for classification of multispectral data, *OP Conference Series: Earth and Environmental Science*, 2014, Vol. 20 (1): 012038. DOI: [10.1088/1755-1315/20/1/012038](https://doi.org/10.1088/1755-1315/20/1/012038).
  24. Friedman J., Hastie T., Tibshirani R. Additive logistic regression: a statistical view of boosting, *Ann. Statist.*, 2000, Vol. 28, № 2, pp. 337–407. DOI: [10.1214/aos/1016218223](https://doi.org/10.1214/aos/1016218223).
  25. Vivek Verma. Application of Bayesian Analysis in Medical Diagnosis, *Journal of the Practice of Cardiovascular Science*, 2019, Vol. 5, P. 136–141. DOI: [10.4103/jpcs.jpcs\\_51\\_19](https://doi.org/10.4103/jpcs.jpcs_51_19).
  26. Alfonso Perez G., Caballero Villarraso J. Alzheimer Identification through DNA Methylation and Artificial Intelligence Techniques, *Mathematics*, 2021, No. 9, 2482. DOI: <https://doi.org/10.3390/math9192482>.

Received 31.10.2022.  
Accepted 22.12.2022.

УДК 519.876.5:681.518.2

## ОБЧИСЛЮВАЛЬНІ ІНТЕЛЕКТУАЛЬНІ МЕТОДИ СТРАТИФІКАЦІЇ ПАЦІЄНТІВ У СИСТЕМАХ МЕДИЧНОГО МОНИТОРИНГУ

**Бакуменко Н. С.** – канд. техн. наук, доцент, доцент кафедри теоретичної та прикладної системотехніки Харківського національного університету імені В.Н. Каразіна, Харків, Україна.

**Стрілець В. Є.** – канд. техн. наук, доцент кафедри теоретичної та прикладної системотехніки Харківського національного університету імені В. Н. Каразіна, Харків, Україна.

**Угрюмов М. Л.** – д-р техн. наук, професор, професор кафедри теоретичної та прикладної системотехніки Харківського національного університету імені В. Н. Каразіна, Харків, Україна.

**Зеленський Р. О.** – асистент кафедри урології та дитячої урології Харківської медичної академії післядипломної освіти, Харків, Україна.

**Угрюмова К. М.** – канд. техн. наук, доцент кафедри математичного моделювання та штучного інтелекту Національного аерокосмічного університету ім. М.Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

**Старенький В. П.** – д-р мед. наук, професор, завідувач відділення радіаційної онкології, ДУ «Інститут медичної радіології та онкології ім. С.П. Григор'єва» Національної академії медичних наук України, Харків, Україна.

**Артюх С. В.** – канд. мед. наук, старший науковий співробітник групи променевої терапії відділу радіології ДУ «Інститут медичної радіології та онкології ім. С.П. Григор'єва» НАМН України, асистент кафедри радіології та радіаційної медицини Харківського національного медичного університету, Харків, Україна.

**Насонова А. М.** – канд. мед. наук, клінічний онколог, завідувачка відділенням клінічної онкології та гематології ДУ «Інститут медичної радіології та онкології ім. С.П. Григор'єва» НАМН України, Харків, Україна.

#### АНОТАЦІЯ

**Актуальність.** У сучасній медичній практиці все більше впроваджується автоматизація й інформаційні технології для діагностування захворювань, моніторингу стану пацієнта, визначення програми лікування тощо. Тому розробка нових і удосконалення існуючих методів стратифікації пацієнтів у системах медичного моніторингу є своєчасною і необхідною.

**Метод.** Розроблено методологію, яка включає такі методи обчислювального інтелекту для стратифікації пацієнтів у системах медичного моніторингу, як:

1) метод кластерного аналізу на основі агентного підходу – визначення можливої кількості станів пацієнтів з використанням контрольованих змінних станів;

2) метод побудови робастних метамоделей за допомогою штучних нейронних мереж при апріорній невизначеності даних (відома лише точність вимірювань) за даними моніторингу стану пацієнтів: а) багатовимірна логістична регресійна модель у вигляді аналітичних залежностей апостеріорних ймовірностей різних станів пацієнтів від контрольованих змінних станів; б) багатовимірна діагностична модель у вигляді аналітичних залежностей цільових функцій (критеріїв якості стану хворого) від контрольованих змінних станів;

3) метод оцінки інформативності контрольованих змінних станів при невизначеності апріорних даних;

4) метод побудови робастних багатовимірних моделей контролю стану пацієнтів при апріорній невизначеності даних у даних моніторингу у вигляді аналітичних залежностей, що прогнозуються за вимірними значеннями контрольованих змінних станів у процесі моніторингу;

5) метод зменшення розмірності простору контрольованих змінних станів на основі аналізу інформативності змінних робастних багатовимірних моделей управління станом пацієнтів (аналіз чутливості);

6) метод визначення станів пацієнтів на основі вирішення задачі класифікації за значеннями контрольних та прогнозованих контрольованих змінних стану з використанням імовірнісних нейронних мереж;

7) метод синтезу раціональної індивідуальної програми лікування хворих у системі медичного моніторингу для стану, виявленого на основі прогнозу.

У роботі запропонована структура моделі вибору раціональної індивідуальної програми лікування пацієнтів на основі IT Data Stream Mining, яка реалізує концепцію «Big Data for Better Outcomes».

**Результати.** Розроблені передові методи обчислювального інтелекту для прогнозування станів використовувалися при виборі тактики лікування пацієнтів, прогнозуванні ускладнень лікування та оцінці виживності пацієнта до та під час спеціального лікування.

**Висновки.** Представлено досвід впровадження концепції «Big Data for Better Outcomes» для вирішення проблеми розробки передових методологій нових стратегій стратифікації пацієнтів. Розроблено передову методологію, методи обчислювального інтелекту для стратифікації пацієнтів у системах медичного моніторингу та прикладну інформаційну технологію, що їх реалізує. Розроблені передові методи прогнозування станів можуть бути використані при виборі тактики лікування хворих, прогнозуванні ускладнень лікування та оцінці виживності хворого до та під час спеціального лікування.

**КЛЮЧОВІ СЛОВА:** інформаційна технологія потокового аналізу даних, системи медичного моніторингу, методи машинного навчання, математичні моделі і методи стратифікації пацієнтів.

УДК 519.876.5:681.518.2

#### ВЫЧИСЛИТЕЛЬНЫЕ ИНТЕЛЛЕКТУАЛЬНЫЕ МЕТОДЫ СТРАТИФИКАЦИИ ПАЦИЕНТОВ В СИСТЕМАХ МЕДИЦИНСКОГО МОНИТОРИНГА

**Бакуменко Н. С.** – канд. техн. наук, доцент, доцент кафедри теоретической и прикладной системотехники Харьковского национального университета имени В. Н. Каразина, Харьков, Украина.

**Стрелец В. Е.** – канд. техн. наук, доцент кафедры теоретической и прикладной системотехники Харьковского национального университета имени В. Н. Каразина, Харьков, Украина.

**Угрюмов М. Л.** – д-р техн. наук, профессор, профессор кафедры теоретической и прикладной системотехники Харьковского национального университета имени В. Н. Каразина, Харьков, Украина.

**Зеленский Р. А.** – ассистент кафедры урологии и детской урологии Харьковской медицинской академии последиplomного образования, Харьков, Украина.

**Угрюмова Е. М.** – канд. техн. наук, доцент кафедры математического моделирования и искусственного интеллекта Национального аэрокосмического университета им. Н.Е. Жуковского «Харьковский авиационный институт», Харьков, Украина.

**Старенький В. П.** – д-р мед. наук, профессор, заведующий отделением радиационной онкологии, ГО «Институт медицинской радиологии и онкологии им. П. С. Григорьева» Национальной академии медицинских наук Украины, Харьков, Украина.

**Артиух С. В.** – канд. мед. наук, старший научный сотрудник группы лучевой терапии отдела радиологии ГО «Институт медицинской радиологии и онкологии им. П. С. Григорьева» Национальной академии медицинских наук Украины, ассистент кафедры радиологии и радиационной медицины, Харьковский национальный медицинский университет, Харьков, Украина.

**Насонова А. Н.** – канд. мед. наук, клинический онколог, заведующая отделением клинической онкологии и гематологии ГО «Институт медицинской радиологии и онкологии им. П. С. Григорьева» Национальной академии медицинских наук Украины, Харьков, Украина.

## АННОТАЦИЯ

**Актуальность.** В современной медицинской практике все больше внедряются автоматизированные и информационные технологии для диагностики заболеваний, мониторинга состояния пациентов, выбора программы лечения и др. В связи с этим разработка новых и усовершенствование существующих методов вычислительного интеллекта для стратификации пациентов в системах медицинского мониторинга является своевременной и необходимой задачей.

**Метод.** Разработана методология, включающая следующие методы вычислительного интеллекта для стратификации пациентов в системах медицинского мониторинга:

1) метод кластерного анализа на основе агентного подхода – определение возможного количества состояний пациентов с использованием контролируемых переменных состояний;

2) метод построения робастных метамоделей с помощью искусственных нейронных сетей при априорной неопределенности данных мониторинга (известна только точность измерений) в данных мониторинга: а) многомерная логистическая регрессионная модель в виде аналитических зависимостей апостериорных вероятностей различных состояний пациентов от контрольных и контролируемых переменных состояний; б) многомерная диагностическая модель в виде аналитических зависимостей целевых функций (критериев качества состояния больного) от контрольных и контролируемых переменных состояний.

3) метод оценки информативности контролируемых переменных состояний при неопределенности априорных данных;

4) метод построения робастных многомерных моделей контроля состояния пациентов при априорной неопределенности данных мониторинга в виде аналитических зависимостей, прогнозируемых по измеренным значениям контрольных и контролируемых переменных состояний в процессе мониторинга;

5) метод уменьшения размерности пространства контролируемых переменных состояний на основе анализа информативности переменных робастных многомерных моделей управления состоянием пациентов;

6) метод определения состояний пациентов на основе решения задачи классификации по значениям контрольных и прогнозируемых контролируемых переменных состояний с использованием вероятностных нейронных сетей;

7) метод синтеза рациональной индивидуальной программы лечения больных в системе медицинского мониторинга, для состояния, выявленного на основе прогноза.

В работе предложена структура модели выбора рациональной индивидуальной программы лечения пациентов на основе IT Data Stream Mining, которая реализует концепцию «Big Data for Better Outcomes».

**Результаты.** Разработанные усовершенствованные методы вычислительного интеллекта для прогноза состояний использовались при выборе тактики лечения пациентов, для прогнозирования осложнений лечения и оценки излечимости пациента до и во время проведения специального лечения.

**Выводы.** Представлен опыт реализации концепции «Большие данные для лучших результатов» для решения задачи разработки передовых методологий для новых стратегий стратификации пациентов. Разработаны передовая методология, методы вычислительного интеллекта для стратификации пациентов в системах медицинского мониторинга и реализующие их прикладные информационные технологии. Разработанные усовершенствованные методы прогноза состояний могут быть использованы при выборе тактики лечения больных, для прогнозирования осложнений лечения и оценки излечимости больного до и во время проведения специального лечения.

**КЛЮЧЕВЫЕ СЛОВА:** информационная технология потокового анализа данных, системы медицинского мониторинга, методы машинного обучения, математические модели и методы стратификации пациентов.

## ЛІТЕРАТУРА

1. Conejar R. J. A Medical Decision Support System (DSS) for Ubiquitous Healthcare Diagnosis System / R. J. Conejar, H.-K. Kim // International Journal of Software Engineering and Its Applications. – 2014. – Vol. 8, № 10. – P. 237–244. DOI: <http://dx.doi.org/10.14257/ijseia.2014.8.10.22>.
2. Waghlikar K. B. Modeling Paradigms for Medical Diagnostic Decision Support: A Survey and Future Directions / K. B. Waghlikar, V. Sundararajan, A. W. Deshpande // Journal of Medical Systems. – 2012. – Vol. 36, Issue 5. – P. 3029–3049. DOI: <https://doi.org/10.1007/s10916-011-9780-4>.
3. Diagnostic model and information technology of classification states in the differential diagnosis NSCLC (non small cell lung cancer) patients with different methods of radiotherapy and chemotherapy / [V. Starenkiy, V. Goryachaya, O. Sokolov, E. Ugryumova] // Journal of Health Sciences. – 2013. – Vol. 3 (8). – P. 7–26. ISSN: 1429-9623.
4. Forecasting of Patients Condition in the Monitoring Medical Systems on the Example of Prostate Diseases / [V. A. Goriacha, O. Yu. Sokolov, K. M. Ugryumova et al.] // Journal of Education, Health and Sport. – 2016. – Vol. 6, № 5. – P. 77–93. DOI: <http://dx.doi.org/10.5281/zenodo.51160>.
5. Bakumenko N. Application of the C-Means Fuzzy Clustering Method for the Patient's State Recognition Problems in the Medicine Monitoring Systems [Electronic resource] / N. Bakumenko, V. Strilets, M. Ugryumov // 3rd International Conference on Computational Linguistics and Intelligent Systems (COLINS-2019), Kharkiv, Ukraine, April 18–19, 2019: pro-

- ceedings. – 2019. – Vol. I: Main Conference. – 10 p. Access mode: <http://ceur-ws.org/Vol-2362/paper20.pdf>
6. Application of artificial neural networks in the problems of the patient's condition diagnosis in medical monitoring systems / [V. Strilets, N. Bakumenko, S. Chernysh et al.] // *Integrated Computer Technologies in Mechanical Engineering – Synergetic Engineering. International Scientific and Technical Conference (ICTM 2019). Kharkov, Ukraine, November, 28–30, 2019.* – Springer Link: 2020. – P. 173–185. DOI: [https://doi.org/10.1007/978-3-030-37618-5\\_16](https://doi.org/10.1007/978-3-030-37618-5_16).
  7. Machine Learning Methods in Medicine Diagnostics Problem / [V. Strilets, N. Bakumenko, V. Donets et al.] // *16th International Conference on ICT in Education, Research and Industrial Applications. Integration, Harmonization and Knowledge Transfer, Kharkiv, Ukraine, 6–10 October 2020: proceedings. – Volume II: Workshops, 2020.* – P. 89–101. Access mode: <https://ceur-ws.org/Vol-2732/20200089.pdf>.
  8. Bakumenko N. Synthesis Method of Robust Neural Network Models of Systems and Processes / [N. Bakumenko, V. Strilets, I. Menailov et al.] // *Integrated Computer Technologies in Mechanical Engineering – 2020 / M. Nechyporuk, V. Pavlikov, D. Kritskiy (eds) – Springer, Cham.* – P. 3–16. DOI: [https://doi.org/10.1007/978-3-030-66717-7\\_1](https://doi.org/10.1007/978-3-030-66717-7_1).
  9. Стратифікація пацієнтів у системах медичного моніторингу : монографія / [В. П. Старенький, С. В. Артюх, О. І. Зеленський та ін.]; под ред. О. Г. Авруніна, С. Б. Радзишевської. – Харків : ХНМУ, 2021. – 148 с.
  10. Radiosensitization by cyclooxygenase-2 inhibitors in radiation therapy of malignant neoplasms: monograph / [N. A. Mitryaeva, V. P. Starenkiy, N. V. Billozor et al.]. – Kharkiv : State Institution “Institute of Medical Radiology and Oncology named SP Grigorieva National Academy of Medical Sciences of Ukraine”, 2021. – 136 p.
  11. A Method for Assessing the Risks of Complications in Chemoradiation Treatment of Squamous Cell Carcinoma of the Head and Neck / [V. Starenkiy, S. Artiukh, M. Ugryumov et al.] // *The Open Bioinformatics Journal.* – 2021. – Vol. 14. – P. 139–143. DOI: [10.2174/18750362021140100138](https://doi.org/10.2174/18750362021140100138).
  12. Agent-Oriented data clustering for medical monitoring / [V. Strilets, V. Donets, M. Ugryumov et al.] // *Radioelectronic and Computer Systemsthis.* 2022. – № 1. – P. 103–114. DOI: <https://doi.org/10.32620/reks.2022.1.08>
  13. Комп'ютерна програма «Методи нелінійного оцінювання в багатокритеріальних задачах робастного оптимального проектування та діагностування систем в умовах параметричної апіорної невизначеності (методологія, методи, методики і реалізують їх комп'ютерні системи підтримки і прийняття рішень)» («ROD&IDS»): Свідectво про реєстрацію авторського права на твір № 82875 / М. Л. Угрюмов, Є. С. Меньяйлов, С. В. Черниш, К. М. Угрюмова (Україна). – Авторське право і суміжні права. Офіційний бюлетень. – К.: Міністерство економічного розвитку і торгівлі України, 2018. – № 51. – С. 403.
  14. Методи машинного навчання у задачах системного аналізу і прийняття рішень: монографія / В. Є. Стрілець, С. І. Шматков, М. Л. Угрюмов та ін.]. – Харків : Харківський національний університет імені В. Н. Каразіна, 2020. – 195 с. – ISBN 978-966-285-627-9.
  15. Sidey-Gibbons J. A. M. Machine Learning in medicine: practical introduction / Jenny A.M. Sidey-Gibbons, Chris J Sidey-Gibbons // *BMC Medical Research Methodology.* – 2019. – No. 19. DOI: [10.1186/s12874-019-0681-4](https://doi.org/10.1186/s12874-019-0681-4).
  16. Varoquaux G. Machine Learning for medical imaging: methodological failures and recommendations for the future / G. Varoquaux, V. Cheplygina // *Digit. Med.* – 2022. – No. 5(48). DOI: <https://doi.org/10.1038/s41746-022-00592-y>.
  17. Antoniou T. Evaluation of machine learning solutions in medicine / T. Antoniou, M. Mamdani // *CMAJ: Canadian medical association Journal.* – 2021. – No. 193(36). – E1425-E1429. DOI: [10.1503/cmaj.210036](https://doi.org/10.1503/cmaj.210036).
  18. Seyhan A. A. Are innovation and new technologies in precision medicine paving a new era in patients centric care? / A. A. Seyhan, C. Carini // *Journal of Translational Medicine.* – 2019. – № 17. – Article 114. DOI: <https://doi.org/10.1186/s12967-019-1864-9>
  19. Nayyar Anand. Chapter 2 – Machine learning in healthcare: review, opportunities and challenges / Nayyar Anand, Gadhavi Lata, Zaman Noor // *Machine Learning and the Internet Medical Things in Healthcare.* – 2021. – P. 23–45. DOI: <https://doi.org/10.1016/B978-0-12-821229-5.00011-2>.
  20. Stratification of the severity of critically ill patients with classification trees / [J. Trujillano, M. Badia, L. Serviá et al.] // *BMC Medical Research Methodology.* – 2009. – Vol. 9, No. 83. DOI: <https://doi.org/10.1186/1471-2288-9-83>.
  21. Amigo J. M. Mathematical methods in medicine: Neuroscience, cardiology and pathology / J. M. Amigo, M. Small // *Philosophical Transactions of The Royal Society A Mathematical Physical and Engineering Sciences.* – 2017. – vol. 375 (2096):20170016. DOI: [10.1098/rsta.2017.0016](https://doi.org/10.1098/rsta.2017.0016).
  22. Azar A. Decision tree classifiers for automated medical diagnosis / A. Azar, S. El-Metwally // *Neural Computing and Applications.* – 2013. – Vol. 23. – P. 2387–2403. DOI: [10.1007/s00521-012-1196-6](https://doi.org/10.1007/s00521-012-1196-6).
  23. Bahari N. I. S. Application of support vector machine for classification of multispectral data / N. I. S. Bahari, A. Ahmad, B. M. Aboobaidar // *OP Conference Series: Earth and Environmental Science.* – 2014. – Vol. 20 (1): 012038. DOI: [10.1088/1755-1315/20/1/012038](https://doi.org/10.1088/1755-1315/20/1/012038)
  24. Friedman J. Additive logistic regression: a statistical view of boosting / J. Friedman, T. Hastie, R. Tibshirani // *Ann. Statist.* – 2000. – Vol. 28, № 2. – P. 337–407. DOI: [10.1214/aos/1016218223](https://doi.org/10.1214/aos/1016218223).
  25. Vivek Verma. Application of Bayesian Analysis in Medical Diagnosis / Vivek Verma // *Journal of the Practice of Cardiovascular Science.* – 2019. – Vol. 5. – P. 136–141. DOI: [10.4103/jpcs.jpcs\\_51\\_19](https://doi.org/10.4103/jpcs.jpcs_51_19).
  26. Alfonso Perez G. Alzheimer Identification through DNA Methylation and Artificial Intelligence Techniques / G. Alfonso Perez, J. Caballero Villarraso // *Mathematics.* – 2021. – № 9. – 2482. DOI: <https://doi.org/10.3390/math9192482>.

## METHOD AND SOFTWARE COMPONENT MODEL FOR SKIN DISEASE DIAGNOSIS

**Lovkin V. M.** – 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.

**Oliinyk A. O.** – Dr. Sc., Professor, Professor of the Department of Software Tools, National University “Zaporizhzhia Polytechnic”, Zaporizhzhia, Ukraine.

**Myronenko N. V.** – Student of the Department of Software Tools, National University “Zaporizhzhia Polytechnic”, Zaporizhzhia, Ukraine.

### ABSTRACT

**Context.** The problem of skin disease diagnosis was investigated in the paper. Its actuality is caused by the necessity of automation of at least advisory medical decision making. Such decisions are made in telemedicine, for instance, when skin disease diagnostics is performed under specific conditions. These conditions are specified by situations when data for analysis are collected but a qualified doctor has no possibility to process the data and to make a diagnosis decision based on it. The object of the study is a process of skin disease diagnosis.

**Objective.** The objective of the study is to develop a skin disease diagnosis method to automate making of advisory medical diagnosis decisions and to increase efficiency of such decisions.

**Method.** The skin disease diagnosis method was proposed in the work. This method applies the modified ResNet50 model. It was proposed to add layers to the ResNet50 model and to train it using transfer learning and fine-tuning techniques. The method also defines image processing in particular through the change of its resolution and uses oversampling technique to prepare a dataset for model training.

**Results.** Experimental investigation of the proposed method was performed using the HAM10000 dataset which contains images of skin diseases. The images were collected using dermatoscopy method. The dataset contains observations for 7 different skin diseases. The proposed method is characterized by the accuracy of 96.31% on this dataset. It is improved accuracy in comparison with the existing neural network models. Software component model was created to give a possibility to integrate the proposed method into a medical diagnosis system.

**Conclusions.** The obtained results of the investigation suggest application of the proposed skin disease method in medical diagnostic system to make advisory decisions by the system and to support making final decisions by a doctor.

**KEYWORDS:** dermatoscopy, medical diagnosis, convolutional neural network, skin disease, ResNet50 model, software component model.

### ABBREVIATIONS

AdaBoost is Adaptive Boosting;  
AI is Artificial intelligence;  
AKIEC is Actinic Keratoses and Intraepithelial Carcinoma;  
ANN is Artificial Neural Network;  
AUC is Area Under Curve;  
BB is Balanced Bagging;  
BCC is Basal Cell Carcinoma;  
BKL is Benign Keratosis-like Lesions;  
BRF is Balanced Random Forest;  
CNN is Convolutional Neural Network;  
DF is Dermatofibroma;  
HAM10000 is Human Against Machine with 10000 training images;  
LR is Logistic Regression;  
MEL is Melanoma;  
NV is Melanocytic Nevi;  
ReLU is Rectified Linear Unit;  
ResNet is Residual Network;  
RF is Random Forest;  
SVM is Support Vector Machine;  
VASC is Vascular Lesions;  
WHO is World Health Organization;

XGBoost is eXtreme Gradient Boosting.

### NOMENCLATURE

$\delta$  is a standard deviation of pixels;  
 $A$  is an accuracy;  
 $C$  is a number of correctly classified observations;  
 $D$  is a set of diagnoses;  
 $d^A$  is a skin disease diagnosis for an image of skin area which was not diagnosed by a dermatologist as it was classified by a model;  
 $d_i$  is an actual skin disease diagnosis for the  $i$ -th observation in a dataset;  
 $d_i^C$  is a skin disease diagnosis for the  $i$ -th observation in a dataset as it was classified by a model;  
 $d_i$  is a class of the  $i$ -th observation in a dataset;  
 $f$  is a functional dependence;  
 $M$  is a number of misclassified observations;  
 $N$  is a number of observations in a dataset;  
 $p_i$  is the  $i$ -th image in a dataset;  
 $p_i^{av}$  is an average pixel intensity value calculated through all values of the  $i$ -th image;  
 $p_i^{g,k}$  is a pixel intensity in all channels for pixel in the  $g$ -th row and  $k$ -th column of the  $i$ -th image;



$pn_i^{q,k}$  is a normalized pixel intensity in all channels for a pixel in the  $q$ -th row and  $k$ -th column of the  $i$ -th image;  
 $S$  is a dataset;  
 $S_{tr}$  is a training dataset;  
 $S_t$  is a test dataset;  
 $S_v$  is a validation dataset;  
 $s^A$  is an image of skin area which was not diagnosed by a dermatologist;  
 $s_i$  is an **imabe** of the  $i$ -th observation in a dataset.

## INTRODUCTION

Despite active current progress of medicine, skin diseases are one of the most dangerous and widespread types of diseases. At the same time cases of skin diseases increase in number. Such trends are caused by either unfavorable external factors (for instance, environmental degradation, climate change, influence of ultraviolet radiation) or individual factors (chronic stress, heredity, unhealthy lifestyle, bad quality of diagnostics or untimely diagnostics). This year the trends are additionally reinforced in Ukraine by the war. Access of displaced persons or Ukrainian citizens in general to quality medicine (specialists with experience in a specified area) is complicated in some territories.

According to the data presented by WHO, the number of cases with detected nonmelanoma skin cancer increased to the level of 1.2 million and the number of melanoma cases reached more than 0.3 million in 2020 [1]. The percentage of fatal cases in these cases is the following: 5% for nonmelanoma skin cancer and 17% for melanoma skin cancer [1].

Official statistics in Ukraine confirm skin cancer as the cancer type with the biggest number of cases: there were more than 180 thousand of active cases in 2020 [2].

Skin diseases are more easily defined in comparison with the other types of cancer in particular. But people do not always consider its symptoms, paying necessary attention in visiting a doctor. Such a way may not only be the reason for serious health problems but also for a death of a human.

The source [3] states that only 40% of nonmelanoma skin tumors are detected in the early stages (the first and the second stages) in Ukraine. It should be noted that early detection guarantees effective treatment of a patient. At the same time melanoma is detected infrequently at the initial stage in Ukraine. But in this case a disease is best treated.

The low quality of diagnostics is caused either by low level of patient information awareness regarding these diseases or by low level of procedure quality. The last one is connected with insufficient qualifications of a doctor, lack of necessary equipment, similarity of disease symptoms, analyses collected with procedural violations, human factor.

Diagnostic process includes gathering of data on patient's condition and drawing a conclusion on health issues. The process is realized through the following stages:  
– detection of health issues by a patient;

– patient's referring to a medical institution for medical aid;  
– collection, integration and interpretation of patient's medical data;  
– generation of hypothesis on illness;  
– confirmation/update of hypothesis via additional obtained data;  
– making final diagnosis [4].

Gathering of patient's medical data for skin diseases requires physical examination of a patient. The examination is traditionally organized using dermatoscopy method. It requires to capture an image of an examined skin area and to estimate the following criteria:

- skin pigmentation change;
- skin color uniformity;
- structure characteristics;
- skin contours and pattern change;
- skin peeling and inflammation;
- skin tightening;
- skin cracks and ulcer areas [5].

Estimations of the presented criteria results in making a diagnosis and determining a treatment.

**The object of the study** is a process of skin disease diagnosis.

Because these diseases have a great impact on the quality of human life and can be dangerous for human health, timely detection of skin neoplasms with increasing accuracy of classification without doctor participation is actual and important problem which should be solved. This classification is necessary in telemedicine support systems, in systems where advisory decisions are made to decrease working time of a doctor with the necessary qualification.

Dermatologist makes his job in a way similar to image recognition by neural network: it demands input image (skin lesions) processed through neural network (skills and knowledge of dermatologist who analyses an image and synthesizes data) to classify it (to make a diagnosis) [6]. In addition, early diagnostics of a disease is one of necessary conditions for successful treatment of a disease, so modern approaches and methods of diagnosis support based on AI (in particular CNN models) should be applied to increase classification accuracy.

In general, neural network models are widely used in the field of dermatology presently taking into account the described similarity between real process and such models. ANN models are assumed to be efficient diagnosis tool [6, 7]. It can be applied to improve recognition results and to decrease number of misdiagnosed cases.

**The subject of the study** are ANN models which are applied for diagnostics of skin diseases based on classification of images which present skin state observations collected using dermatoscopy method.

**The objective of the study** is to develop a skin disease diagnosis method to automate making of advisory medical diagnosis decisions and to increase efficiency of such decisions.

## 1 PROBLEM STATEMENT

Skin disease diagnosis problem lies in a recognition of skin area images using appropriate mathematical model. In this work the problem should be specified using CNN as a mathematical model.

Let  $S$  be a dataset presenting skin disease cases with known diagnoses.

The dataset  $S = \{s_i\}$ ,  $i = 1, 2, \dots, N$  consists of  $N$  observations  $s_i$ . Each  $i$ -th observation consists of a pair  $s_i = \langle p_i, d_i \rangle$ , where  $p_i$  is an object which presents image of skin area and  $d_i$  is a diagnosis which was made by dermatologist based on an image  $p_i$ .

Object  $p_i$  is an array  $\{p_i^{q,k}\}$  of elements. Each element  $p_i^{q,k}$  is a record of intensity for pixel in the  $q$ -th row and  $k$ -th column. This record presents pixel intensity for each channel of an image.

Let  $D$  be a set of diagnoses. A set  $D$  contains the following elements:

- $d^0$  – Actinic Keratoses and Intraepithelial Carcinoma (to represent investigation results AKIEC is used to mark this diagnosis);
- $d^1$  – Basal Cell Carcinoma (marked by BCC in investigation results);
- $d^2$  – Benign Keratosis-like Lesions (marked by BKL in investigation results);
- $d^3$  – Dermatofibroma (marked by DF in investigation results);
- $d^4$  – Melanocytic Nevi (marked by NV in investigation results);
- $d^5$  – Melanoma (marked by MEL in investigation results);
- $d^6$  – Vascular Lesions (marked by VASC in investigation results).

In this study the diagnoses from  $d_0$  to  $d_6$  are used in a set  $D$ . This set could be extended if it is necessary in the medical diagnostic system and there is appropriate dataset.

The skin disease diagnosis problem is to determine a functional dependence  $f(s_i) = d_i^C$  which classifies an image  $s_i$  into a diagnosis  $d_i^C \in D$ . This functional dependence should define a model  $f$  minimizing accuracy  $A$ .

Accuracy  $A$  is calculated as a ratio of correctly classified  $C$  observations to misclassified  $M$ :  $A = C / M \cdot 100$ . Number  $C$  is calculated as a total number of correctly classified observations between  $N$  observations of a dataset. Number  $M$  is calculated as a total number of misclassified observations between  $N$  observations of a dataset.

The defined model  $f$  should be used to classify new diagnosis case  $s^A$  into a class  $d^A \in D$ .

## 2 REVIEW OF THE LITERATURE

AI-based approach to medical diagnostics in general and to diagnostics of skin diseases in particular became widespread recently [8–9]. In a few years a lot of AI-based diagnosis methods were investigated. To develop an own method to support one of the stages of medical diagnostic system according to the objective of the study it is necessary to analyze existing solutions in the beginning.

In the paper [10] different classifiers, including SVM, LR, RF, AdaBoost, BB, BRF, were investigated for melanoma detection using images collected by dermatoscopy method. The best results were obtained by BRF method. SVM, LR and RF were unstable. AdaBoost was a stable classifier but with the lowest AUC value between all the considered methods. AUC was used because the problem in the paper was considered as a binary classification problem. Classification was performed for two classes: melanoma and benign tumor. The main shortcoming of the study is in the binary classification. At the same time this study allowed to compare different classifiers.

The methods investigated in the paper [10] are still applied for images recognition in medical diagnosis but ANN became the most actively used model for solving these problems. ANN popularity is conditioned on its higher productivity, adaptability to different types of objects and its ability to increase classification accuracy [11].

In the paper [12] the method for diagnostics based on the proposed model of ANN was applied for skin disease image recognition. Machine learning methods, including RF, SVM and XGBoost methods, VGG16 and ResNet models, were investigated for the problem solving. As a result of this investigation ANN had significantly better recognition results in comparison with the other considered machine learning methods. The main conclusion of this investigation is in the efficiency and applicability of ANN for the problem solving. The model proposed in the paper [12] was characterized by less estimation results in comparison with VGG16 and ResNet models. The ResNet model had the highest accuracy (90.5%) in the investigation.

The paper [13] was aimed at investigation of different models of CNN, including VGG16, VGG19, MobileNet and InceptionV3 with transfer learning, for recognition of skin disease images. Main feature of the study is in the modification of these models by extended convolution technique, where filter is applied to an area which is bigger than filter length by excluding input values with a step, which is equivalent to a convolution with big filter, created from initial filter by placing zeros in it [14]. The highest value of recognition accuracy was provided by InceptionV3 (90.95%) model. The obtained accuracy value should be increased for practical application in medical diagnostic system.

In the paper [15] efficiency of ANN application for skin disease recognition was investigated in comparison with diagnostics by general practitioners. The dataset was processed for usage of two variants of images: images with high and low resolution. VGG16, VGG19, ResNet34, ResNet50, ResNet101, SEResNet50, EfficientNetB5, MobileNet models were investigated. The highest accuracy (83.88% for images with high resolution and 82.47% for images with low resolution) was obtained using MobileNet model. The main result of this investigation is in the higher recognition accuracy of the worst investigated model in comparison with general practitioners. It means that only decisions made by experienced specialists are valuable. At the same time the

following decisions could be made using computer systems:

- advisory decisions;
- decisions in situations when there is no possibility of access to a specialist;
- preliminary decisions on what specialist is needed and if this specialist is needed.

The investigated works have such shortcomings as application of outdated methods, classification into less number of classes (for example, the work [10]) and besides low recognition accuracy which is a major indicator for the subject area. These drawbacks are a basis for the next study of skin disease diagnosis methods and models to improve the known results.

### 3 MATERIALS AND METHODS

To achieve the objective of the study the skin disease diagnosis method was proposed. It consists of the following stages.

At the first stage of the method data presenting images of skin diseases from a set  $D$  collected using dermatoscopy method should be processed to prepare it for a dataset  $S$  creation. This preparation procedure includes change of image resolution to guarantee higher speed of model training because of the less size of images in dataset (investigated and confirmed in the study [16]). Higher training speed is important for practical application of the method in medical diagnosis system because it decreases time necessary for update when new data presenting new cases are included to the dataset and model retraining is performed.

At the second stage a dataset  $S$  ready for training should be created.

During this stage dataset  $S$  should be separated into the following subdatasets: training dataset  $S_r$  and validation dataset  $S_v$ . During this study of the method test dataset  $S_t$  should be separated from dataset  $S$ . Then these subdatasets should be normalized using a procedure which calculates a value of each pixel. Each pixel  $p_i^{q,k}$  should be transformed into  $pn_i^{q,k}$  by subtracting average value  $p_i^{av}$  calculated through all values and by dividing this result by standard pixel deviation  $\delta$ :  $pn_i^{q,k} = (p_i^{q,k} - p_i^{av}) / \delta$ .

A dataset is possibly unbalanced, because a number of observations referred to one class may significantly exceed a number of observations referred to another classes. Such an issue notably impacts on model recognition accuracy, because classes, which observations are in minority, may be ignored by a model. It results in misdiagnosing. Oversampling technique should be used to solve this issue. But oversampling has its own practical drawback. After its application images will be duplicated in training dataset for classes, causing model overtraining. It is explained by image learning without its analysis and detection of specified details of each disease. As a result this model will not be able to highlight structural patterns of tumors and could not be applied for new data observations, which were not used for training. Such methods of image transformation as rotation, horizontal and vertical

© Lovkin V. M., Subbotin S. A., Oliinyk A. O., Myronenko N. V., 2023  
 DOI 10.15588/1607-3274-2023-1-4

shift, wheeling and scaling should be used to solve the issue.

At the third stage a model, which defines functional dependence  $f$ , should be created by modification of pre-trained ResNet50 model. This model was chosen, because it fits the problem solved in the study, using connections with quick access, missing one layer or some layers avoiding model overtraining and increasing diagnosis accuracy [17].

The structure of the full modified model is shown in Fig. 1.

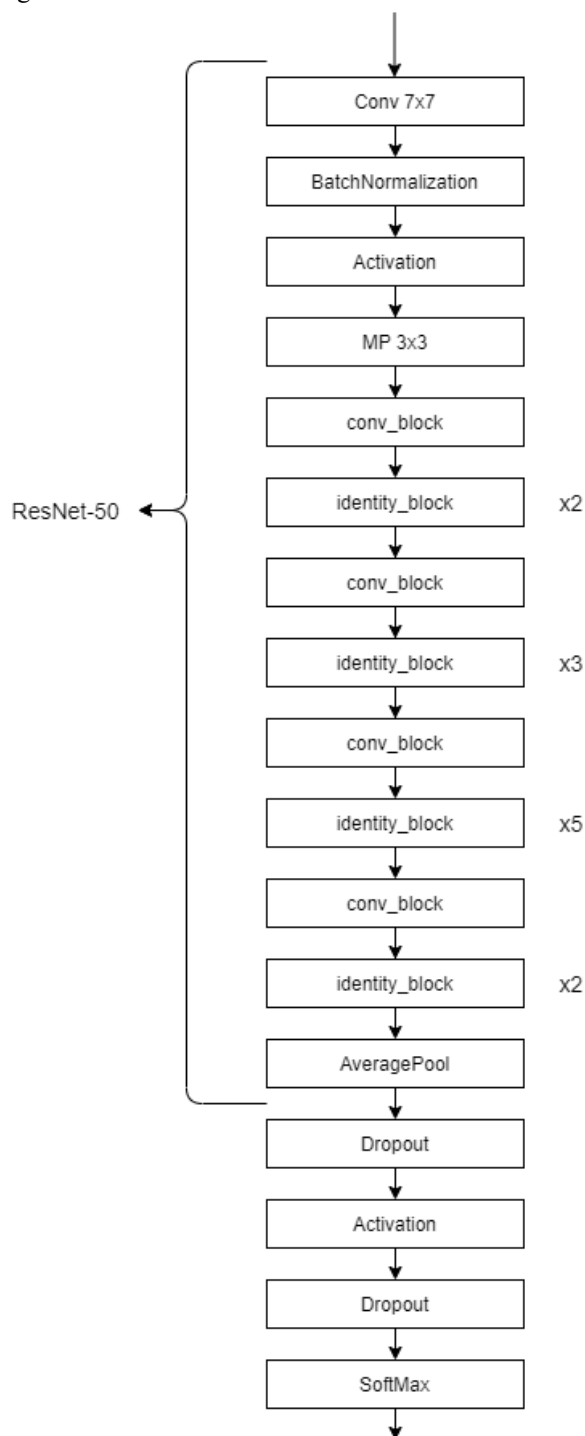


Figure 1 – Structure of the modified model

The structure of the ResNet50 model was extended by the following additional layers:

- dropout layer;
- layer with 128 neurons and ReLu activation function;
- dropout layer;
- layer with 7 neurons and SoftMax activation function.

The structure of identity block of ResNet50 model is shown in Fig. 2.

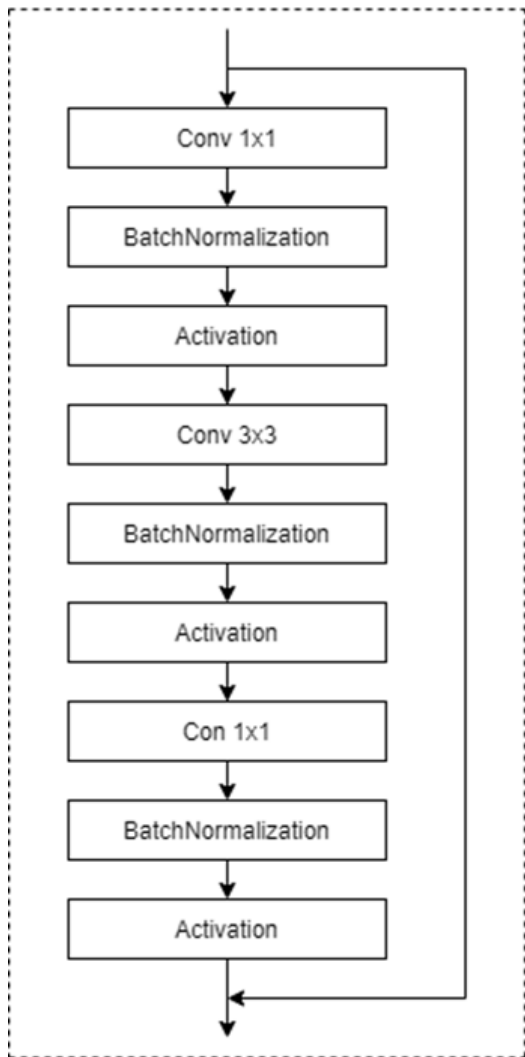


Figure 2 – Structure of identity block of ResNet50 model

The structure of convolutional block of ResNet50 model is shown in Fig. 3.

At the fourth stage the created model is trained. Transfer learning [18] and fine-tuning [19] should be used for training. Transfer learning allows to reuse pre-trained model for new problem. Fine-tuning allows to fix weights of first layers, changing only last layers (aiming at more specific forms and objects on an image) of the model during training.

At the next stages the trained model is applied for new cases when skin diseases should be diagnosed. Each new case should be investigated through the following stages:

- image creation using dermatoscopy method;
- image preprocessing through size change, transformation to a multidimensional array, normalization of its values;
- image recognition, including transfer of preprocessed array to the trained model, classification by the model, return of a class number  $d^A \in D$  and of the corresponding diagnosis.

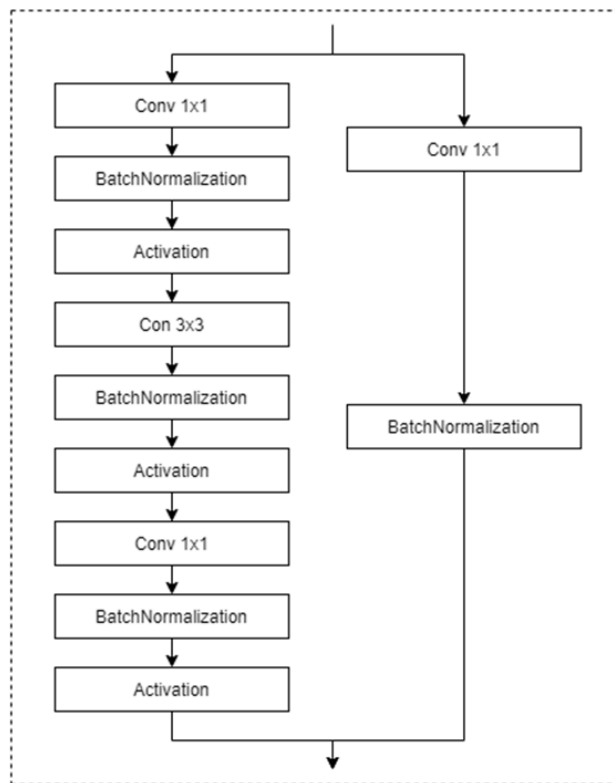


Figure 3 – Structure of convolutional block of ResNet50 model

#### 4 EXPERIMENTS

Experimental investigation was held using the HAM10000 dataset [20]. The dataset contains 10015 images of skin diseases. The images were collected using dermatoscopy method. The dataset contains observations for 7 different skin diseases (AKIEC, BCC, BKL, DF, NV, MEL, VASC).

Histopathology (53% cases), further examination of a patient (37% cases), consensus of experts (9%) and microscopy (1%) were used to detect diagnosis for every case in the dataset. Localization of observations of skin diseases is presented in Fig. 4.

This dataset was separated into subdatasets in the following manner: 60% of data were used as training dataset  $S_{tr}$ , 20% were used as validation dataset  $S_v$  and 20% were used as test dataset  $S_t$ .

Preliminary analysis of the HAM10000 dataset defined that the biggest number of cases in the dataset represents NV (6705). It is 6 times more than number of MEL (1113) and BKL (1099) cases. Four other diagnoses are represented by even less number of cases. Fraction of every diagnosis in the initial dataset is visualized in Fig. 5.

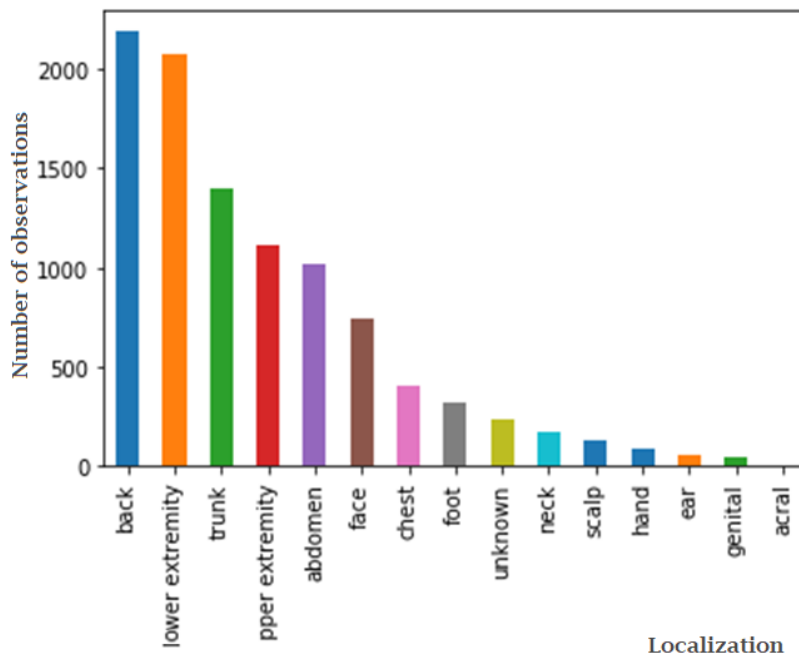


Figure 4 – Graph of localization of observations in the dataset

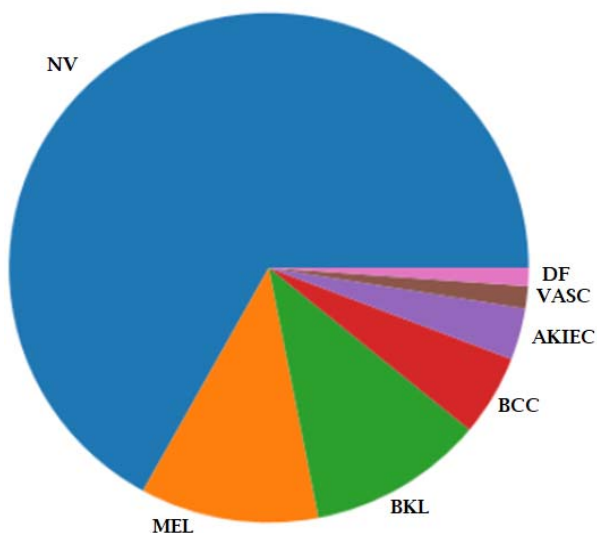


Figure 5 – Part of disease cases in the initial dataset

Oversampling technique was applied to balance number of cases in the processed dataset. Number of cases for each diagnosis is presented in Table 1.

Table 1 – Number of images in the dataset used for investigation after oversampling

Diagnosis	Number of images
AKIEC	4905
BCC	5140
BKL	5495
DF	5750
MEL	5565
NV	6705
VASC	5680

Experimental investigation was performed using the processed dataset. The following models were investigated:

- VGG16 in modifications used in the papers [12], [15], [21];
- VGG19 [15];
- ResNet in modifications used in the papers [12], [21];
- ResNet34 [15];
- ResNet101 [15];
- SEResNet50 [15];
- ResNeXt101 [22];
- AlexNet [21];
- DenseNet [21];
- MobileNet in modifications in the papers [15], [21];
- EfficientNetB5 [15];
- Ensemble [22];
- the proposed model using appropriate method.

The proposed model was trained during 60 epochs with Adam optimizer, which is used instead of stochastic gradient descent algorithm and updates weights in network iteratively, making possible to change learning rate for each parameter and to adapt weight values according to the learning process [23]. Dropout value was set to 0.5. Last 22 layers of the model were used for training, because fine-tuning technique was applied in the method, providing only last layers training.

Experimental investigation was performed using software which was developed to support method practical application. It was developed using the following tools:

- TensorFlow and Keras libraries for model creation and training;
- NumPy, Pandas, scikit learn libraries for data processing;
- Python programming language;

– Matplotlib library for data visualization.

Additional training parameters included: mini-batch size – 16, minimum learning rate – 0.00001, maximum learning rate – 0.001, reduce learning rate function – ReduceLROnPlateau, loss function – categorical\_crossentropy.

Software component model was created to give a possibility to integrate the proposed method into a medical diagnosis system. Software component model provides software interface to communicate with class model. The interface includes methods for the following tasks:

- to create modified model;
- to upload a dataset;
- to preprocess uploaded dataset;
- to train a model;
- to save trained model into a file;
- to upload saved model from a file;
- to upload new case;
- to preprocess new diagnosis case;
- to classify preprocessed diagnosis case using trained model.

This decision allows to integrate this component model into medical diagnosis system in the future. Medical diagnosis system makes decisions for different tasks using accumulated patient data. This approach allows to add necessary components after this system is developed and to create system after separate component models are created if these component models use the same approach.

## 5 RESULTS

During training of the proposed model values of recognition accuracy (Fig. 6) and error function (Fig. 7) presented by categorical crossentropy were visualized in dynamics.

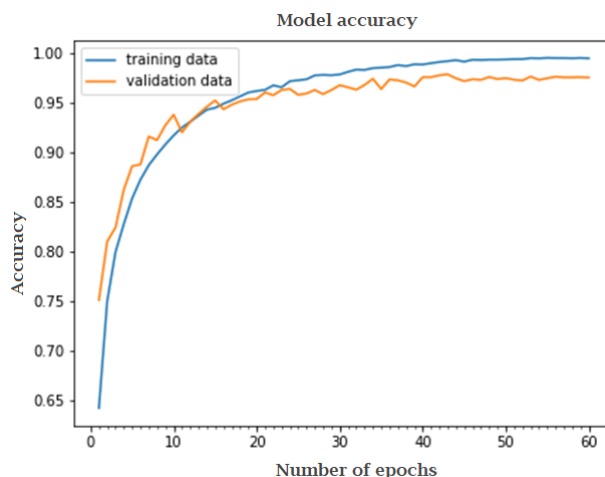


Figure 6 – Graph of accuracy change during model training for training and validation datasets

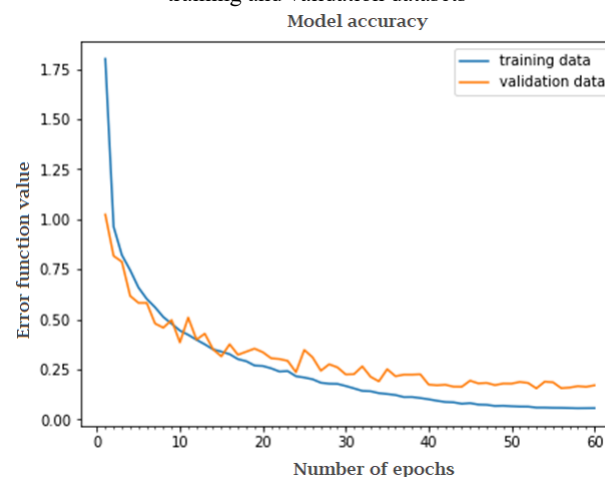


Figure 7 – Graph of error function value change during model training for training and validation datasets

Appropriate values of accuracy and error function were fixed in every epoch for training and validation datasets (Fig. 8).

```
Epoch 1/60
1570/1569 [=====] - 104s 66ms/step - loss: 1.8020 - acc: 0.6424 - val_loss: 1.0240 - val_acc: 0.7512
Epoch 2/60
1570/1569 [=====] - 94s 60ms/step - loss: 0.9633 - acc: 0.7502 - val_loss: 0.8172 - val_acc: 0.8102
Epoch 3/60
1570/1569 [=====] - 95s 60ms/step - loss: 0.8218 - acc: 0.7996 - val_loss: 0.7881 - val_acc: 0.8243
Epoch 4/60
1570/1569 [=====] - 91s 58ms/step - loss: 0.7460 - acc: 0.8279 - val_loss: 0.6170 - val_acc: 0.8624
Epoch 5/60
1570/1569 [=====] - 94s 60ms/step - loss: 0.6595 - acc: 0.8535 - val_loss: 0.5827 - val_acc: 0.8861
.....
Epoch 55/60
1570/1569 [=====] - 87s 55ms/step - loss: 0.0581 - acc: 0.9950 - val_loss: 0.1858 - val_acc: 0.9745
Epoch 56/60
1570/1569 [=====] - 86s 55ms/step - loss: 0.0581 - acc: 0.9948 - val_loss: 0.1564 - val_acc: 0.9761
Epoch 57/60
1570/1569 [=====] - 88s 56ms/step - loss: 0.0569 - acc: 0.9949 - val_loss: 0.1593 - val_acc: 0.9755
Epoch 58/60
1570/1569 [=====] - 88s 56ms/step - loss: 0.0559 - acc: 0.9947 - val_loss: 0.1663 - val_acc: 0.9755
Epoch 59/60
1570/1569 [=====] - 87s 55ms/step - loss: 0.0563 - acc: 0.9949 - val_loss: 0.1630 - val_acc: 0.9756
Epoch 60/60
1570/1569 [=====] - 87s 55ms/step - loss: 0.0566 - acc: 0.9946 - val_loss: 0.1706 - val_acc: 0.9753
```

Figure 8 – Model training results

Confusion matrix (Table 2) was created to assess recognition quality for every skin disease from the set  $D$ . This matrix represents number of images which were true classified and misclassified by the proposed model. A row in the matrix represents actual values, and a column represents results of classification by the proposed model.

Table 2 – Confusion matrix for skin disease recognition using the proposed method

Diagnosis	AKIEC	BCC	BKL	DF	NV	MEL	VASC
AKIEC	66	0	0	0	0	0	0
BCC	0	105	0	0	1	0	0
BKL	0	0	196	2	1	0	0
DF	0	0	0	22	0	0	0
NV	0	4	22	23	1286	13	1
MEL	0	0	0	1	5	222	0
VASC	0	0	0	1	0	0	32

Fraction of true classified images for the proposed model is presented in Fig. 9. The results are presented for each diagnosis separately.

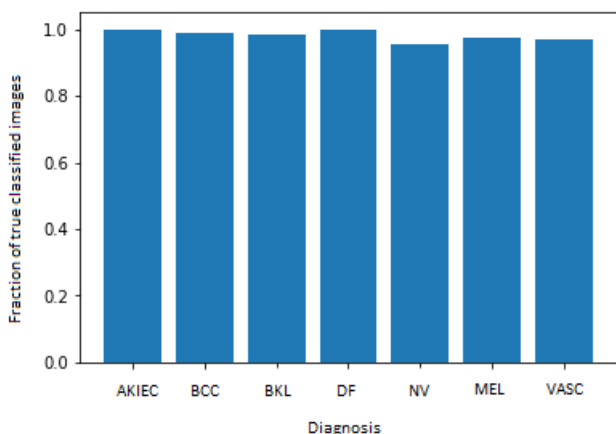


Figure 9 – Chart displaying fraction of true classified images for every diagnosis

Appropriate fraction of misclassified images for the proposed model is presented in Fig. 10.

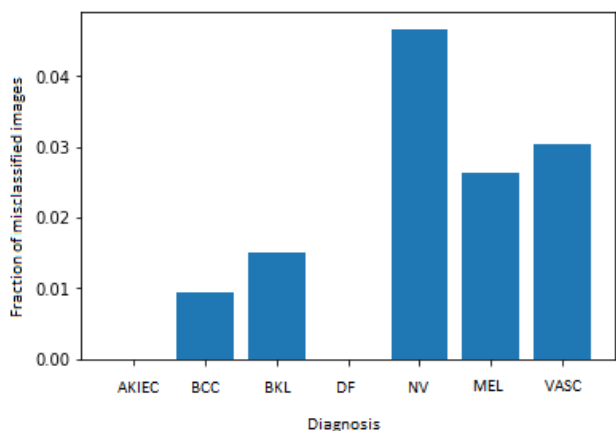


Figure 10 – Chart displaying fraction of misclassified images for every diagnosis

All observations, which actual skin disease diagnosis  $d_i$  from the dataset  $S_i$ , is not equal to a value  $d_i^c$  classified by the model, are considered as misclassified.

Accuracy of the proposed and other investigated models was estimated using the test dataset. Results of accuracy comparison are presented in Table 3.

Table 3 – Results of classification by different CNN models

Model	Source	Accuracy A, %
VGG16	[12]	78.00
ResNet	[12]	90.50
AlexNet	[21]	88.81
ResNet	[21]	85.20
VGG16	[21]	86.09
DenseNet	[21]	85.25
MobileNet	[21]	82.62
ResNet34	[15]	78.84
ResNet101	[15]	78.96
SEResNet50	[15]	80.72
VGG16	[15]	81.65
VGG19	[15]	79.02
EfficientNetB5	[15]	77.14
MobileNet	[15]	84.73
ResNeXt101	[22]	93.20
Ensemble	[22]	92.83
The modified ResNet50 model of the proposed method	–	96.31

## 6 DISCUSSION

A wide investigation (Table 3) of different CNN models was conducted to clarify its practical applicability for skin disease diagnosis problem solving. For the purpose HAM10000 dataset was used. The dataset includes cases for 7 skin diseases. Since other investigations were mainly conducted for a smaller number of skin disease such an investigation has practical significance.

The results of experimental investigation confirm that propositions for modification of mathematical model based on ResNet50, presented in the appropriate method, allowed to increase classification accuracy. The modified model has accuracy of 96.31% on test dataset, which is 5.81% higher than base ResNet model and which is 3.11% higher than the best model in a group of other investigated models (ResNeXt101). The best results were obtained by ResNet (90.5%), Ensemble (92.83%), ResNeXt101 (93.2%) and the modified ResNet50 model. Other models had accuracy lower than 90%.

Accuracy of models with less number of layers was lower: VGG16, VGG19, ResNet34, MobileNet models all had accuracy of less than 90% and closer to 80%. It approves that such a number of layers is not enough to learn all patterns necessary to recognize all diseases which were analyzed.

The proposed model allowed to get the number of misclassified observations which was not higher than 5% for all diagnoses (Fig. 10). All AKIEC and DF cases were true classified (Table 2). NV had 1349 observations in test dataset. The proposed model misclassified 63 observations: 4 cases were diagnosed as BCC, 22 cases as BKL, 23 cases as DF, 13 cases as MEL and 1 case as VASC.

This diagnosis is characterized by the lowest accuracy of 95.33%.

The modified ResNet50 model has additional layers, in such a way having large demands on dataset necessary for model training. Oversampling technique together with transformation methods, applied to dataset images, allowed to get a wider training dataset with sufficient number of observations in it. Images used as input for the model had size of 100x75 pixels. Learning curves (Fig. 6 and Fig. 7) together with model training results (Fig. 8) approve this statement, demonstrating that there was no under- or overtraining. The modified model has accuracy of 97.53% on validation dataset

These experimental results allow to recommend to integrate the proposed method into medical diagnosis system for practical usage as an advisory block [24]. It might have significant effect in countries or territories, where equal and wide access to quality medicine and qualified specialists is impossible or temporarily restricted.

### CONCLUSIONS

The problem of skin disease diagnosis was investigated in the paper. It was considered as classification problem. The existing methods for the problem solving are not perfect and may be improved. These methods classify observations into less number of classes and have lower accuracy. The skin disease diagnosis method, based on the modified ResNet50 model, was proposed.

To increase model recognition accuracy it was proposed to perform preliminary data processing to balance dataset, to reduce image size, to normalize values, to transform images inside dataset using different methods to avoid oversampling.

ResNet50 model was modified adding new layers. Dropout layer, layer with 128 neurons and ReLu activation function, dropout layer, and layer with 7 neurons and SoftMax activation function were added to the model. The extended model was trained using transfer learning and fine-tuning techniques Main learning settings were defined. Separate stages were defined in the method to apply the trained model to diagnose new cases.

Experimental investigation was performed using the HAM10000 dataset, which consists of 10015 images of 7 skin diseases (AKIEC, BCC, BKL, DF, NV, MEL, VASC). Accuracy of the proposed model and the existing models was compared during experimental investigation. The proposed method with the modified ResNet50 model applied for classification had an accuracy 96.31 %, which was the highest result between all investigated models. This accuracy level of the method together with learning results confirms possibility of its practical implementation.

Software component model was created to provide an interface for integration of the proposed method into a medical diagnosis system. This interface provides means to manage classification model and to use it for decision making.

**The scientific novelty** of the obtained results is in the proposed method of skin disease diagnosis.

**The practical significance** of the obtained results is in classification accuracy of the proposed method and in software, which was developed to support the proposed method, as well as software component model, which allows to integrate the developed software into medical diagnosis system in the future.

**Prospects for the further research** are to integrate the proposed method and software, which was created based on appropriate software component, for skin disease diagnosis into medical diagnosis system.

### ACKNOWLEDGEMENTS

The work was performed as part of the state budget research projects “Development of methods and tools for analysis and prediction of dynamic behavior of nonlinear objects” (state registration number 0121U107499) of Software Tools Department of National University “Zaporizhzhia Polytechnic”.

### REFERENCES

1. Sung H. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries, *Statistics*, 2021, Volume 71, pp. 209–249. DOI: 10.3322/caac.21660.
2. Rak v Ukraine 2020. Informaciya k razmyshleniyu – analiz danykh kancer-reestra [Electronic resource]. 2021. Access mode: <https://chemoteka.com.ua/blog/article/rak-v-ukraine-2020-informaciya-k-razmyshleniyu-analiz-danykh-kancer-reestra-197>. [In Russian]
3. Rak shkiry (pukhlyny shkiry) [Electronic resource]. Access mode: <https://spizhenko.clinic/uk/zabolevaniya/opuholizkozhi-uk>. [In Ukrainian]
4. Balogh E. P., Miller B. T., Ball J. R. Improving Diagnosis in Health Care, Washington, DC, National Academies Press, 2015, 472 p. DOI: 10.17226/21794.
5. Dermatoskopiya [Electronic resource]. Access mode: <https://guideh.clinic/dermatologiya/dermatoskopiya/>. [In Russian]
6. Convolutional neural networks in dermatology [Electronic resource]. 2018. Access mode: <https://dermnetnz.org/topics/convolutional-neural-networks-in-dermatology>.
7. Minagawa A., Koga H., Sano T. Dermoscopic diagnostic performance of Japanese dermatologists for skin tumors differs by patient origin: A deep learning convolutional neural network closes the gap, *The Journal of Dermatology*, 2020, Volume 48, pp. 232–236. DOI: 10.1111/1346-8138.15640.
8. Pham T. C., Tran G. S., Nghiem T. P., Doucet A., Luong C. M., Dung H. V. A Comparative Study for Classification of Skin Cancer, 2019 International Conference on System Science and Engineering (ICSSE), Jul 2019, Dong Hoi, Vietnam, pp. 267–272. DOI:10.1109/ICSSE.2019.8823124.
9. Pratt M. K. Artificial intelligence in primary care, *Medical Economics Journal*, 2018, Volume 95, Issue 15, pp. 19–30.
10. Artificial Intelligence in Medical Diagnosis [Electronic resource]. 2021. Access mode: <https://sma.org/ai-in-medical-diagnosis/>.
11. Yadav S. S., Jadhav S. M. Deep convolutional neural network based medical image classification for disease diagnosis, *Journal of Big Data*, 2019, Volume 6, pp. 113–131. DOI: 10.1186/s40537-019-0276-2.



12. Garg R., Maheshwari S., Shukla A. Decision Support System for Detection and Classification of Skin Cancer Using CNN, *Innovations in Computational Intelligence and Computer Vision*, 2021, Volume 1, pp. 578–586. DOI: 10.48550/arXiv.1912.03798.
13. Rab Ratul M. A., Mozaffari M. H., Lee W.-S., Parimbelli E. Skin Lesions Classification Using Deep Learning Based on Dilated Convolution [Electronic resource], bioRxiv, 2019, pp. 1–11. Access mode: <https://www.biorxiv.org/content/10.1101/860700v3.full.pdf>. DOI : 10.1101/860700.
14. Dilated Convolution [Electronic resource]. 2022. Access mode: <https://www.geeksforgeeks.org/dilated-convolution/>.
15. Lucius M., De All J., De All J. A., Belvisi M., Radizza L., Lanfranconi M., Lorenzatti V., Galmarini C. M. Deep Neural Frameworks Improve the Accuracy of General Practitioners in the Classification of Pigmented Skin Lesions, *Diagnostics*, 2020, Volume 10, pp. 1–15. DOI: 10.3390/diagnostics10110969.
16. Nelson J. You Might Be Resizing Your Images Incorrectly [Electronic resource]. 2020. Access mode: <https://blog.roboflow.com/you-might-be-resizing-your-images-incorrectly/>.
17. Feng V. An Overview of ResNet and its Variants [Electronic resource]. 2017. Access mode: <https://towardsdatascience.com/an-overview-of-resnet-and-its-variants-5281e2f56035>.
18. Donges N. What Is Transfer Learning? Exploring the Popular Deep Learning Approach [Electronic resource]. Access mode: <https://builtin.com/data-science/transfer-learning>.
19. Fine-tuning Convolutional Neural Network on own data using Keras Tensorflow [Electronic resource]. Access mode: <https://cv-tricks.com/keras/fine-tuning-tensorflow/>.
20. Tschandl Ph. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions [Electronic resource]. 2018. Access mode: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DBW86T>. DOI: 10.7910/DVN/DBW86T.
21. Ali S., Miah S., Haque J. An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models, *Machine Learning with Applications*, 2021, Volume 5, pp. 1–8. DOI: 10.1016/j.mlwa.2021.100036.
22. Chaturvedi S., Tembhurne J., Diwan T. A multi-class skin Cancer classification using deep convolutional neural networks, *Computer Science*, 2020, Volume 79, Issue 39/40, pp. 28477–28498. DOI:10.1007/s11042-020-09388-2.
23. Brownlee J. Gentle Introduction to the Adam Optimization Algorithm for Deep Learning [Electronic resource]. 2021. Access mode: <https://machinelearningmastery.com/adam-optimization-algorithm-for-deep-learning/>.
24. Lovkin V., Oliinyk A., Lukashenko Y. Air Pollution Prediction as a Source for Decision Making Framework in Medical Diagnosis [Electronic resource]. *IntelITSIS'2021: 2nd International Workshop on Intelligent Information Technologies and Systems of Information Security*, 2021, Khmelnytskyi, Ukraine, 2021, pp. 295–302.

Received 20.11.2022.  
Accepted 08.12.2022.

УДК 004.896

## МЕТОД ТА ПРОГРАМНА КОМПОНЕНТНА МОДЕЛЬ ДІАГНОСТУВАННЯ ШКІРНИХ ЗАХВОРЮВАНЬ

**Львкін В. М.** – канд. техн. наук, доцент, 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.

**Олійник А. О.** – д-р техн. наук, професор, професор кафедри програмних засобів Національного університету «Запорізька політехніка», Запоріжжя, Україна.

**Мироненко Н. В.** – студентка кафедри програмних засобів Національного університету «Запорізька політехніка», Запоріжжя, Україна.

### АНОТАЦІЯ

**Актуальність.** У даній статті розглянуто проблему діагностування шкірних захворювань. Актуальність проблеми пояснюється необхідністю автоматизації щонайменше дорадчих рішень в медицині, зокрема при підтримці телемедицини під час діагностування шкірних захворювань, коли можливо отримати придатні для аналізу дані, але лікар, здатний прийняти рішення, не може з тих чи інших причин обробити ці дані власноруч. Об'єктом роботи є процес діагностування шкірних захворювань.

**Мета роботи** – розробити метод діагностування шкірних захворювань для автоматизації дорадчих діагностуючих рішень в медицині і підвищення ефективності таких рішень.

**Метод.** У статті запропоновано метод діагностування шкірних захворювань на основі побудови і навчання моделі ResNet50, модифікованої шляхом додавання нових шарів і використання технік трансферного навчання та fine-tuning. Метод виконує попередню обробку зображень зокрема шляхом зміни їх розміру та використання техніки oversampling для підготовки вибірки даних для навчання моделі.

**Результати.** Експериментальне дослідження проводилося на наборі даних про шкірні захворювання HAM10000, який складається з 10015 зображень захворювань шкіри, отриманих методом дерматоскопії, для 7 різних діагнозів. У результаті застосування модифікованої моделі на основі запропонованого методу точність розпізнавання склала 96.31%, що перевищує точність розглянутих подібних моделей нейронних мереж. Програмна компонента модель була створена для надання можливості інтегрувати запропонований метод до складу системи медичного діагностування.

**Висновки.** Отримані в результаті проведеного дослідження результати дозволяють рекомендувати запропонований метод діагностування шкірних захворювань для впровадження у складі системи медичного діагностування задля забезпечення підтримки винесення дорадчих рішень системою з подальшим винесенням остаточних рішень лікарем.

**КЛЮЧОВІ СЛОВА:** дерматоскопія, медичне діагностування, згортова нейронна мережа, шкірне захворювання, модель ResNet50, програмна компонента модель.

#### ЛІТЕРАТУРА

1. Sung, H. Global Cancer Statistics 2020 : GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries [Text] / H. Sung, J. Ferlay, R. L. Siegel // *Statistics*. – 2021. – Volume 71. – P. 209–249. DOI : 10.3322/саас.21660.
2. Рак в Україні 2020. Інформація к размышлению – анализ данных канцер-реестра [Электронный ресурс]. – 2021. – Режим доступа : <https://chemoteka.com.ua/blog/article/rak-v-ukraine-2020-informaciya-k-razmyshleniyu-analiz-dannyh-kanцер-reestra-197>.
3. Рак шкіри (пухлини шкіри) [Електронний ресурс]. – Режим доступу : <https://spizhenko.clinic/uk/zabolevaniya/opuholi-kozhi-uk>.
4. Balogh, E. P. Improving Diagnosis in Health Care [Text] / E. P. Balogh, B. T. Miller, J. R. Ball. – Washington, DC : National Academies Press, 2015. – 472 p. DOI : 10.17226/21794.
5. Дерматоскопія [Электронный ресурс] – Режим доступа : <https://guideh.clinic/dermatologiya/dermatoskopiya/>.
6. Convolutional neural networks in dermatology [Electronic resource]. – 2018. – Access mode : <https://dermnetnz.org/topics/convolutional-neural-networks-in-dermatology>.
7. Minagawa, A. Dermoscopic diagnostic performance of Japanese dermatologists for skin tumors differs by patient origin : A deep learning convolutional neural network closes the gap [Text] / A. Minagawa, H. Koga, T. Sano // *The Journal of Dermatology*. – 2020. – Volume 48. – P. 232–236. DOI : 10.1111/1346-8138.15640.
8. A Comparative Study for Classification of Skin Cancer [Text] / Tri Cong Pham, Giang Son Tran, Thi Phuong Nghiem, Antoine Doucet, Chi Mai Luong, et al. // 2019 International Conference on System Science and Engineering (ICSSE), Jul 2019, Dong Hoi, Vietnam. – Vietnam : P. 267–272. DOI : 10.1109/ICSSE.2019.8823124.
9. Mary K. Pratt. Artificial intelligence in primary care [Text] / Mary K. Pratt // *Medical Economics Journal*. – 2018. – Volume 95, Issue 15. – P. 19–30.
10. Artificial Intelligence in Medical Diagnosis [Electronic resource]. – 2021. – Access mode : <https://sma.org/ai-in-medical-diagnosis/>.
11. Yadav S. S. Deep convolutional neural network based medical image classification for disease diagnosis [Text] / S. S. Yadav, S. M. Jadhav // *Journal of Big Data*. – 2019. – Volume 6. – P. 113–131. DOI : 10.1186/s40537-019-0276-2.
12. Garg, R. Decision Support System for Detection and Classification of Skin Cancer Using CNN [Text] / R. Garg, S. Maheshwari, A. Shukla // *Innovations in Computational Intelligence and Computer Vision*. – 2021. – Volume 1. – P. 578–586. DOI : 10.48550/arXiv.1912.03798.
13. Skin Lesions Classification Using Deep Learning Based on Dilated Convolution [Electronic resource] / [Md. Aminur Rab Ratul, M. Hamed Mozaffari, W.-S. Lee, E. Parimbelli] // bioRxiv. – 2019. – P. 1–11. – Access mode : <https://www.biorxiv.org/content/10.1101/860700v3.full.pdf>. DOI : 10.1101/860700.
14. Dilated Convolution [Electronic resource]. – 2022. – Access mode : <https://www.geeksforgeeks.org/dilated-convolution/>.
15. Deep Neural Frameworks Improve the Accuracy of General Practitioners in the Classification of Pigmented Skin Lesions [Text] / M. Lucius, J. De All, J. Antonio De All et al. // *Diagnostics*. – 2020. – Volume 10. – P. 1–15. DOI : 10.3390/diagnostics10110969.
16. Nelson J. You Might Be Resizing Your Images Incorrectly [Electronic resource] / J. Nelson. – 2020. – Access mode : <https://blog.roboflow.com/you-might-be-resizing-your-images-incorrectly/>.
17. Feng, V. An Overview of ResNet and its Variants [Electronic resource] / V. Feng. – 2017. – Access mode : <https://towardsdatascience.com/an-overview-of-resnet-and-its-variants-5281e2f56035>.
18. Donges N. What Is Transfer Learning? Exploring the Popular Deep Learning Approach [Electronic resource] / N. Donges. – Access mode : <https://builtin.com/datascience/transfer-learning>.
19. Fine-tuning Convolutional Neural Network on own data using Keras Tensorflow [Electronic resource]. – Access mode : <https://cv-tricks.com/keras/fine-tuning-tensorflow/>.
20. Tschandl, Ph. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions [Electronic resource] / Ph. Tschandl. – 2018. – Access mode : <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DBW86T>. DOI : 10.7910/DVN/DBW86T.
21. Ali S. An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models [Text] / S. Ali, S. Miah, J. Haque // *Machine Learning with Applications*. – 2021. – Volume 5. – P. 1–8. DOI : 10.1016/j.mlwa.2021.100036.
22. Chaturvedi, S. A multi-class skin Cancer classification using deep convolutional neural networks [Text] / S. Chaturvedi, J. Tembhurne, T. Diwan // *Computer Science*. – 2020. – Volume 79, Issue 39/40. – P. 28477–28498. DOI : 10.1007/s11042-020-09388-2.
23. Brownlee, J. Gentle Introduction to the Adam Optimization Algorithm for Deep Learning [Electronic resource] / J. Brownlee. – 2021. – Access mode : <https://machinelearningmastery.com/adam-optimization-algorithm-for-deep-learning/>.
24. Lovkin, V. Air Pollution Prediction as a Source for Decision Making Framework in Medical Diagnosis [Electronic resource] / V. Lovkin, A. Oliinyk, Y. Lukashenko // *IntellT-SIS'2021: 2nd International Workshop on Intelligent Information Technologies and Systems of Information Security*, March 24–26, 2021, Khmelnytskyi, Ukraine. – Khmelnytskyi : KhNU, 2021. – P. 295–302.

## IMAGE SEGMENTATION WITH A CONVOLUTIONAL NEURAL NETWORK WITHOUT POOLING LAYERS IN DERMATOLOGICAL DISEASE DIAGNOSTICS SYSTEMS

**Polyakova M. V.** – Doctor of Technical Sciences, Associate Professor, Professor of the Department of Applied Mathematics and Information Technologies, National University “Odessa Polytechnic”, Odessa, Ukraine.

### ABSTRACT

**Context.** The problem of automating of the segmentation of spectral-statistical texture images is considered. The object of research is image processing in dermatological disease diagnostic systems.

**Objective.** The aim of the research is to improve the segmentation performance of color images of psoriasis lesions by elaboration of a deep learning convolutional neural network without pooling layers.

**Method.** The convolutional neural network is proposed to process a three-channel psoriasis image with a specified size. The initial color images were scaled to the specified size and then inputted on the neural network. The architecture of the proposed neural network consists of four convolutional layers with batch normalization layers and ReLU activation function. Feature maps from the output of these layers were inputted to the 1×1 convolutional layer with the Softmax activation function. The resulting feature maps were inputted to the image pixel classification layer. When segmenting images, convolutional and pooling layers extract the features of image fragments, and fully connected layers classify the resulting feature vectors, forming a partition of the image into homogeneous segments. The segmentation features are evaluated as a result of network training using ground-truth images which segmented by an expert. Such features are robust to noise and distortion in images. The combination of segmentation results at different scales is determined by the network architecture. Pooling layers were not included in the architecture of the proposed convolutional neural network since they reduce the size of feature maps compared to the size of the original image and can decrease the segmentation performance of small psoriasis lesions and psoriasis lesions of complex shape.

**Results.** The proposed convolutional neural network has been implemented in software and researched for solving the problem of psoriasis images segmentation.

**Conclusions.** The use of the proposed convolutional neural network made it possible to enhance the segmentation performance of plaque and guttate psoriasis images, especially at the edges of the lesions. Prospects for further research are to study the performance of the proposed CNN then abrupt changes in color and illumination, blurring, as well as the complex background areas are present on dermatological images, for example, containing clothes or fragments of the interior. It is advisable to use the proposed CNN in other problems of color image processing to segment statistical or spectral-statistical texture regions on a uniform or textured background.

**KEYWORDS:** psoriasis image, image segmentation, convolutional network, pooling layer, color space, deep learning.

### ABBREVIATIONS

BSA is the Body Surface Area;  
PASI is the Psoriasis Area and Severity Index;  
CNN is a convolutional neural network;  
ReLU is a rectified linear unit;  
TP is a number of true positive samples in percent;  
TN is a number of true negative samples in percent;  
FP is a number of false positive samples in percent;  
FN is a number of false negative samples in percent;  
FOM is a Figure of Merit.

### NOMENCLATURE

$I_l$  is the number of pixels of the psoriasis lesion edges, obtained by the proposed CNN;

$I_A$  is the number of pixels of the psoriasis lesion edges on the ground-truth image;

$\alpha$  is a scale factor;

$d_i$  is the distance between the pixel of the psoriasis lesion edge, obtained by the proposed CNN, and the pixel of the psoriasis lesion edge on the ground-truth image, measured along the normal to this edge.

### INTRODUCTION

Medical diagnostic systems are widely used in the field of health care. In particular, such systems can provide information about the pathologies based on medical images. For example, various datasets related to the treatment of dermatological diseases have been collected. A number of datasets include images obtained during the diagnosis of psoriasis. The World Health Organization statistics estimate the number of patients suffering from this disease at 125 million people. According to the International Federation of Psoriasis Associations, the prevalence of this disease depends on the region and ranges from 1.2% to 5%, averaging 3% of the general population [1].

The disease of psoriasis is accompanied by unpleasant sensations such as peeling and itching of the affected areas of the skin, causing discomfort to the patient, and is also a systemic disease, often accompanied by arthritis, diabetes, cardiovascular diseases, etc. To diagnose and monitor the effectiveness of treatment, psoriasis lesions are isolated on photographs of the patient's skin and evaluate their geometric characteristics. The values of these characteristics are used in assessing the clinical severity of psoriasis using the BSA and the PASI [2].

The identification of psoriasis lesions on skin images by a dermatologist requires a significant time and effort. In addition, since there are other skin diseases similar to psoriasis, then the diagnosis of psoriasis also requires significant experience of a specialist in the field of dermatology. Therefore, medical diagnostic systems are used to perform automated processing of images of psoriasis lesions [3–5].

The object of research is image processing in dermatological disease diagnostic systems.

To assess the clinical severity of psoriasis using the BSA and PASI indexes in dermatological disease diagnostic systems, the results of segmentation of skin images are needed to identify the psoriasis areas [4]. Segmentation of color images of skin affected by psoriasis is a difficult problem due to uneven or insufficient lighting, irregular shape and different sizes of psoriasis lesions, blurred boundaries between lesions and normal skin. In addition, the result of segmentation of images of psoriasis lesions is affected by skin pigmentation and texture that varies from person to person, the presence of hair, as well as the type of skin (dry, oily or combination), which determines its reflectivity.

The subject of the research is the methods of segmentation of color images of psoriasis lesions.

Methods for segmenting color images of psoriasis lesions that do not use convolutional neural networks (CNNs) are generally characterized by low performance [6–8], which is due to the natural variability of human skin and the size of psoriasis lesions. So, plaque psoriasis is often characterized by the lesions of considerable size, while guttate psoriasis is appeared by small lesions. The psoriasis image segmentation is also significantly affected by the experience of the researcher and the segmentation feature selection. Therefore, deep learning CNNs have recently been used to segment color images of psoriasis lesions, providing automatic feature selection [9–10]. The most common architecture of such networks includes convolutional layers, pooling layers and fully connected layers. When segmenting images, convolutional and pooling layers extract the features of image fragments, and fully connected layers classify the resulting feature vectors, forming a partition of the image into homogeneous segments. The pooling layers of the CNN are reduce the size of image feature maps, but negatively affects on the segmentation performance, especially for the images of the small psoriasis lesions and complex-shaped psoriasis lesions.

The aim of the research is to improve the of segmentation performance of color images of psoriasis lesions by elaboration of a deep learning CNN without pooling layers.

## 1 PROBLEM STATEMENT

The color image of psoriasis lesions is represented by the vector function  $\mathbf{I}(x,y)=(I_R(x,y), I_G(x,y), I_B(x,y))$ , where  $x=1, \dots, n$ ;  $y=1, \dots, m$ ;  $n, m$  are positive integers,  $n$  is the number of rows,  $m$  is the number of columns of the  
© Polyakova M. V., 2023  
DOI 10.15588/1607-3274-2023-1-5

image. The intensities of the color components  $I_R(x,y)$ ,  $I_G(x,y)$ ,  $I_B(x,y)$  take values from the interval  $[0, 1]$ . Then each pixel of the image with coordinates  $(x,y)$  is described by three features, namely intensities of color components with values  $I_R(x,y)$ ,  $I_G(x,y)$ ,  $I_B(x,y)$ . To identify the psoriasis lesions on the image, it is necessary to segment it into non-overlapping regions corresponding to regions of lesions and normal skin. To do this, each pixel of the original image must be associated with the value of the target feature. There is a label of one of two classes, specifically, 0 for normal skin, 1 for psoriasis lesion. The values of the target feature for the psoriasis image should be represented as a binary image which is the result of segmentation.

To segment the color images of psoriasis lesions, a CNN elaboration is needed. The CNN input is a three-dimensional matrix of intensity values of color components for image pixels. At the output of this network, a two-dimensional matrix of target feature values is obtained for each image pixel. To construct a convolutional neural network  $CNN = \{struct, param\}$ , it is necessary to select its architecture *struct* and evaluate network parameters *param* so as to minimize the cost function on the training set [11, 12]. It is assumed that the training set includes psoriasis images and corresponding ground-truth images for CNN parameters learning.

## 2 REVIEW OF THE LITERATURE

An analysis of the literature shows that psoriasis image processing implies either classification of such images to determine the severity of the disease [13–15] or psoriasis image segmentation [16–18], since the area of psoriatic lesions is taken into account when determining the severity of the disease and evaluating the effectiveness of treatment [2, 4, 10]. When selecting and evaluating the psoriasis image features it is important to consider that normal skin regions differ from psoriasis lesions both in color and texture. For example these regions can be distinguished by the values of the amplitude and frequency of the components of the spectral texture [19, 20]. Therefore, single-scale and multi-scale methods are presented in the literature for psoriasis image processing.

Single-scale methods use color features and/or spectral features at the dominant frequency. Such methods segmenting psoriasis image into areas of normal skin and psoriasis lesions, based on the feature vectors calculated in the neighborhood of each pixel. The the resulting segmentation performance depends on the parameter values setting. In addition there are difficulties in segmenting objects of different sizes and setting of parameters.

Thus, in [7], two color spaces were used to segment psoriatic plaques on skin images, namely CIE Luv and CIE Lch. In the CIE Luv space, the color feature vectors of the pixels were clustered based on the Gaussian mixture model, and in the CIE Lch color space, a semi-wrapped Gaussian mixture model was proposed to solve this problem. For the localization of plaques on the skin, the von Mises distribution was assumed to determine the

confidence intervals for the distribution parameters of normal skin and plaque color. This approach showed a higher segmentation performance compared to the fuzzy c-means, the segmentation accuracy was 79.53%.

In [8], a method for texture segmentation of plaque psoriasis images was elaborated for express diagnostic systems in dermatology. Psoriasis lesions on the images were modelled as fragments of a quasi-periodic texture on a complex background with additive Gaussian noise. To segmenting psoriasis lesions, there was used sequentially the localization of spatial frequencies of the image; the amplitude detecting that transforms the texture features into intensity; the edge detection of the resulting image. Then 87% of image pixels were correctly segmented by this method.

In [6], for the segmentation of color psoriasis images the RGB images convert to the Lab space, and then k-means clustering was applied to the color feature vectors of image pixels. To improve the segmentation performance the merged group of pixels was filtered, as a result of which correctly segmented pixels averaged 95%.

In [19, 20], for systems of express diagnostics of psoriasis, skin images were represented by a spectral-statistical texture. The vector-difference method was applied for image segmentation. This method is based on the calculation of distances between feature vectors of image pixel neighborhoods. Each feature vector is involved the mean value and the standard deviation of the intensities of the neighborhood pixels. For images of plaque psoriasis, the percentage of correct pixel segmentation was 92.7%, and for pixels of lesions of plaque psoriasis, up to 96%, depending on the signal-to-noise ratio.

Multiscale methods are elaborated for segmenting the lesions of different sizes on psoriasis images. However, there are difficulties when combining the results of segmentation at different scales. For example, in [21], a method for segmenting of the psoriasis images in the Lab color space was proposed. In this space, multi-scale superpixels of images were identified, and superpixels that did not correspond to the skin areas were discarded. Next, at each scale, the superpixels were k-means clustered. The merging of segmentation results at different scales was performed using voting. Correctly segmented pixels in the images were about 90%. The method has shown its effectiveness in the treatment of skin with hair, and with lesions of different sizes and shapes.

A separate group of multiscale methods are CNNs that are used both for segmentation and for classifying of psoriasis images.

In [3], a classification system for skin diseases was developed, implemented as Android application based on the MobilNet. This network alternates convolutional layers and depthwise separable convolutions with depthwise and pointwise layers. Each of these 27 layers was followed by batch normalization and a ReLU activation function. The average percentage of correct recognition of the seven skin diseases reached 94.4% after

oversampling to balance the data set and pre-process the researched images.

In [13], images of psoriasis lesions were classified into images of plaque psoriasis and guttate psoriasis. A CNN was used with two convolutional layers, followed by pooling layers, then two fully connected layers were applied. Correct image recognition of plaque and guttate psoriasis was 82.9% and 72.4%.

In [4], psoriasis images were analyzed to automatically form the severity of the disease. The severity of the psoriasis are assessed on the basis of the erythema or redness level, the degree of scaliness, the thickness or induration of psoriasis plaques on the human skin. The percentage of the body surface area affected by psoriasis also included in PASI estimation. Psoriasis images were classified by a CNN that consisted of a shared subnet followed by three parallel subnets. At the output of these subnets, the level of the disease severity was obtained for each of the three features as an integer from 0 to 4. The shared subnet consisted of 5 convolutional layers with the ReLU activation function and normalization after the 1st and 2nd layers, as well as three pooling layers after the 1st, 2nd, 5th convolutional layers. Subnets for each feature included 2 fully connected layers with the ReLU activation function, after which the SoftMax activation function was applied. The correct recognition of the severity scores based on the each feature was 60%, and about 94% if assuming a maximum deviation  $\pm 1$  from the actual scores for any of erythema, scaling and induration scores. Under this assumption, the correct recognition of the severity scores is 85.3%.

In [9], a modified U-Net architecture referred as PsLSNet was proposed for the segmentation of psoriasis images. It is a deep learning CNN with 29 layers. By reducing the covariant shift through the implementation of PsLSNet, the training time is reduced compared to U-Net. The PsLSNet network is able to segment psoriasis images even under poor acquisition conditions and in the presence of artifacts. The average percentage of correctly segmented pixels was 94.8% with a sensitivity of 89.6% and a specificity of 97.6%.

In [10], it was assumed that psoriasis images may be blurred or contain areas of a complex background. For the segmentation of such images, a YOLACT CNN has been developed, which includes four consecutive subnets. These are a subnet for evaluating segmentation features, a subnet for generating feature maps at different image scales, a subnet for classifying the obtained feature vectors, and a subnet for generating the result of segmentation and post-processing. Correct segmentation of image pixels was achieved in 96–97% of cases.

The analysis of methods of psoriasis image segmentation showed that the main characteristic of these methods is the segmentation performance, which is affected by the shape and sizes of isolated lesions that differ for plaque and guttate psoriasis. In the case of express diagnostics, the time of image processing is also significant.

Single-scale methods are characterized by high speed, but the segmentation performance is low, especially when identifying small psoriasis lesions and complex-shaped psoriasis lesions. The results of image segmentation often contain defects in the form of small regions at the boundaries of areas corresponding to psoriasis lesions. In addition, parameters setting are time consuming for single-scale methods.

Multiscale methods are characterized by a higher performance of psoriasis image segmentation due to the better localization of the boundaries of psoriasis lesions. However, these methods are often unable to segment the small lesions or fragments of lesions of complex shape, mistaking them for areas of healthy skin. If the multiscale methods of psoriasis image segmentation don't use the CNN, then it is necessary for the researcher to select segmentation features, the procedure for their evaluation, and a method for combining segmentation results at different scales. The use of CNN allows to avoid these difficulties, since segmentation features are evaluated as a result of network training using ground-truth images which segmented by an expert. Such features are robust to noise and distortion in images. The combination of segmentation results at different scales is determined by the network architecture. However, deep learning CNN contains a large number of layers and, accordingly, is characterized by a significant number of parameters. The training of such CNN requires a large training set, and the complexity of the architecture increases the network training time. In addition, a common drawback of the CNN in solving the problem of psoriasis image segmentation is the low performance if objects of small sizes or complex shapes are processed. This due to the

inclusion of pooling layers in the network architecture, which reduce the scale of image feature maps.

### 3 MATERIALS AND METHODS

The proposed neural network is elaborated to process a three-channel image with a size of 224×224 pixels. Therefore, the initial color images were scaled to the specified size and then giving them as input of the neural network. The architecture of the proposed neural network is shown in Table 1. This CNN contains five convolutional layers, four batch normalization layers, four ReLU activation function layers, one Softmax activation function layer, and one image pixel classification layer. The first convolutional layer uses 18 convolution kernels of size 85×85×3 pixels with a stride of 1 pixel. The feature maps at the output of this layer are batch normalized and then the ReLU activation function is applied. Next, feature maps from the output of the ReLU activation function layer, were inputed to the second convolutional layer, which uses 18 convolution kernels of size 55×55×3 pixels with a stride of 1 pixel. The feature maps at the output of this layer are also batch normalized and then the ReLU activation function is applied again. The resulting feature maps were inputed to the third convolutional layer, which uses 18 convolution kernels of size 35×35×3 pixels with a stride of 1 pixel. The feature maps at the output of this layer are again batch normalized, and then the ReLU activation function is applied again. It is followed by the fourth convolutional layer, which uses 18 convolution kernels of size 15×15×3 pixels with a stride of 1 pixel. The obtained feature maps at the output of this layer are batch normalized, and then the ReLU activation function is applied again. Feature

Table 1 – The proposed CNN architecture

Layer number	Type	Comment	Activations	Learnables
1	Image Input	256×256×3 images with zero center normalization	224×224×3	–
2	Convolution	18 85×85×3 convolutions with stride [1 1] and same padding	224×224×18	Weights: 85×85×3×18 Bias: 1×1×18
3	Batch normalization	Batch normalization with 18 channels	224×224×18	Offset: 1×1×18 Scale: 1×1×18
4	ReLU	Activation function	224×224×18	–
5	Convolution	18 55×55×3 convolutions with stride [1 1] and same padding	224×224×18	Weights: 55×55×3×18 Bias: 1×1×18
6	Batch normalization	Batch normalization with 18 channels	224×224×18	Offset: 1×1×18 Scale: 1×1×18
7	ReLU	Activation function	224×224×18	–
8	Convolution	18 35×35×3 convolutions with stride [1 1] and same padding	224×224×18	Weights: 35×35×3×18 Bias: 1×1×18
9	Batch normalization	Batch normalization with 18 channels	224×224×18	Offset: 1×1×18 Scale: 1×1×18
10	ReLU	Activation function	224×224×18	–
11	Convolution	18 15×15×3 convolutions with stride [1 1] and same padding	224×224×18	Weights: 15×15×3×18 Bias: 1×1×18
12	Batch normalization	Batch normalization with 18 channels	224×224×18	Offset: 1×1×18 Scale: 1×1×18
13	ReLU	Activation function	224×224×18	–
14	Convolution	2 1×1×18 convolutions with stride [1 1] and same padding	224×224×2	Weights: 1×1×18×2 Bias: 1×1×2
15	Softmax	Activation function	224×224×2	–
16	Pixel classification	–	–	–

maps from the output of this layer were inputted to the fifth convolutional layer, which uses 2 convolution kernels of size  $1 \times 1 \times 18$  pixels with a stride of 1 pixel, then the Softmax activation function is applied. The resulting feature maps were applied to the image pixel classification layer.

In the convolution layer, the matrix of the convolution kernel moves through a two-dimensional array of image pixels. The values of the corresponding image elements and the convolution kernel are multiplied the results are added up and inputted on the next layer. Usually, several convolution kernels are used in the convolution layer. In addition, it is the stride parameter, which indicates the convolution kernel shift along the image matrix.

Batch normalization layer solves vanishing gradient problem and train deep neural networks consisting of a few dozen layers. It is known that the error backpropagation algorithm converges faster if the input data is normalized (has zero mean and unit variance) [13]. However, when a signal propagates through a neural network, its mean value and variance can change significantly. To avoid this, the standard normalization of the outputs of the convolutional layer is performed by subtracting from each output  $x_i$  the average value  $\mu_i$  of the input packet examples by dimension  $i$  and dividing by  $(\sigma_i + \varepsilon)^{0.5}$ , where  $\sigma_i$  is the standard deviation of the input packet images by dimension  $i$ ,  $\varepsilon$  is a constant that ensures the stability of calculations. However, normalizing the output of a neural network convolutional layer can change the representation of the data in the next layer. Therefore, two parameters are introduced: compression  $\gamma_i$  and shift  $\beta_i$  of the normalized value for each output, which are adjusted in the learning process along with the rest of the parameters and transform the normalized output value  $x_{ni}$  as  $y_i = \gamma_i x_{ni} + \beta_i$ . For convolutional neural networks, batch normalization reduces training time and reduces the chance of overfitting.

The ReLU activation function  $\text{ReLU}(y) = \max(0, y)$  returns 0 for a negative argument, and in the case of a positive argument, returns the same. The advantages of this function over the sigmoid are the fast calculation of the derivative (for negative arguments it is 0, for positive ones it is 1), and the sparseness of activation (fewer neurons being activated).

Pooling layers were not included in the architecture of the proposed CNN since they reduce the size of feature maps compared to the size of the original image and can decrease the segmentation performance of small psoriasis lesions and psoriasis lesions of complex shape.

#### 4 EXPERIMENTS

As a result of the experiment, the segmentation performance was estimated for 50 images of psoriasis lesions from the [22]. The researched images contained only areas of healthy skin and psoriasis lesions. Areas corresponding to the background were absent. The image sizes varied as follows. For the images of plaque psoriasis the height varied from 200 to 380 pixels, and the width

varied from 180 to 500 pixels. The height of the images of guttate psoriasis varied from 400 to 800 pixels, the width from 300 to 530 pixels. These images were marked by an expert on areas of normal skin and psoriasis lesions. Then the segmentation performance was estimated by confusion matrices [23, 24]. Since psoriasis images were segmented on health skin and psoriasis lesions, then the confusion matrix of size  $2 \times 2$  consists of the following elements. TP is the percentage of image pixels from the class labeled "Normal skin" that are correctly assigned to the class labeled "Normal skin"; FP is the percentage of image pixels from the class labeled "Psoriasis lesion" that are incorrectly assigned to the class labeled "Normal skin"; FN is the percentage of image pixels from the class labeled "Normal skin" incorrectly assigned to the class labeled "Psoriasis lesion"; TN is the percentage of image pixels from the class labeled "Psoriasis lesion" that are correctly assigned to the class labeled "Psoriasis lesion".

To characterize the detection of the edges of psoriasis lesions, the FOM value [25] was used:

$$FOM = \frac{1}{I} \sum_{i=1}^{I_A} \frac{1}{1 + \alpha d_i^2},$$

where  $I = \max(I_I, I_A)$ . The FOM value is normalized such that  $FOM=1$  for a well detected edge. The factor  $1/I$  characterizes doubled and split edges.

When training the proposed CNN, a cross-entropy loss function was used, for which the relative frequencies  $w$  and  $1-w$  of the appearance of pixels of the "Normal skin" and "Psoriasis lesion" classes were taken into account. For each image pixel, the value of cross-entropy  $L$  was calculated by the formula

$$L = -(1/w) t_1 \log_2 y_1 - 1/(1-w) t_2 \log_2 y_2,$$

where  $t_i = 1$  if the image pixel is assigned to class  $i$  on the ground-truth image, otherwise  $t_i = 0$ , with  $i = 1$  for the class "Normal skin" and  $i = 2$  for the class "Psoriasis lesion". The  $y_i$  value is the result of calculating the value of the Softmax function for the image pixel under consideration. It is interpreted as the probability that an image pixel belongs to the class "Normal skin" ( $i = 1$ ) or "Psoriasis lesion" ( $i = 2$ ). To calculate the cross-entropy of the image, the values of  $L$  were summed over all its pixels.

First, as part of the experiment, RGB, Lab, YCbCr, HSV, YIQ, XYZ color spaces were used to represent images of psoriasis lesions [26]. In this part of experiment the proposed CNN and U-Net [27] consisting of four levels realized by  $3 \times 3$  convolutions are researched. The stochastic gradient descent with a moment of 0.9 and an initial learning rate of 0.01 was used to train the proposed CNN. The Adam method with an initial learning rate of 0.0001 was used to train the U-Net.

Note that the color images of lesions of plaque and guttate psoriasis differ significantly in their properties, primarily in the size of the lesions, as well as their color

and texture. Plaque psoriasis usually presents by red or light gray lesions of considerable size with coarse grained texture, while guttate psoriasis presents by small reddish lesions of fine texture. Therefore, in this research, segmentation performance was evaluated separately for images of plaque and guttate psoriasis, in contrast to the papers [6–10, 19–21] and others, in which segmentation performance measures were averaged over images of both classes.

Second, as part of the experiment, the Adam method with an initial learning rate of 0.05 and the stochastic gradient descent with a moment of 0.9, and an initial learning rate of 0.01 was used to train the proposed CNN on the researched psoriasis images. To avoid overfitting of the network, training was stopped when the value of the loss function began to increase. The number of training epochs depended on the color space in which the images were presented.

Third, as part of the experiment, the psoriasis image segmentation performance for the proposed CNN was compared with the known segmentation methods [6–10, 19–21]. The accuracy measure is used. It represents the integral characteristic of the segmentation performance calculated by averaging of *TP* and *TN*.

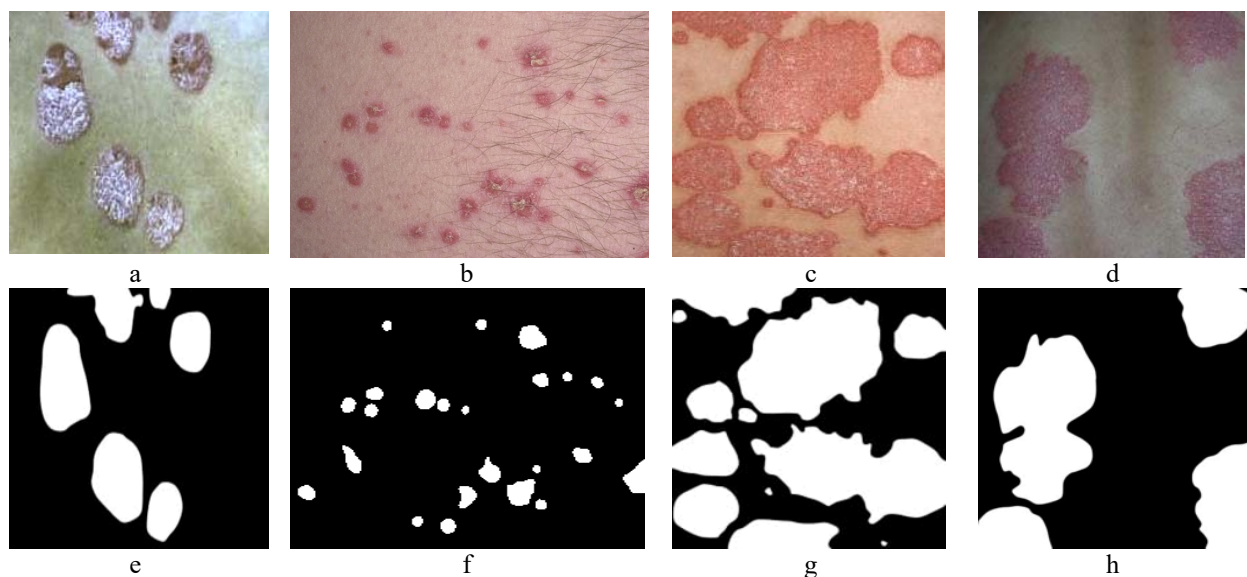
And at last, the comparison of processing time was made for the image segmentation by the proposed CNN and the U-Net [27] for the researched psoriasis images scaled to a size of 224×224 pixels.

## 5 RESULTS

The elements of confusion matrices for the results of psoriasis image segmentation with representation in different color spaces is shown in Table 2. Values in this table were obtained by averaging the *FOM*, *TP*, and *TN* for the segmentation of plaque psoriasis images, guttate psoriasis images, and all the researched psoriasis images.

Table 2 – The values of *TP* and *TN* for the results of segmentation by the proposed CNN and U-Net of psoriasis images represented in different color spaces

Color space	Plaque psoriasis images			Guttate psoriasis images			All researched psoriasis images		
	<i>TR</i>	<i>TN</i>	<i>FOM</i>	<i>TR</i>	<i>TN</i>	<i>FOM</i>	<i>TR</i>	<i>TN</i>	<i>FOM</i>
Proposed CNN									
RGB	98.48	99.55	0.7844	93.41	91.83	0.7637	95.53	96.78	0.7741
Lab	98.37	99.72	0.7256	91.57	88.01	0.5102	94.42	95.51	0.6179
HSV	97.92	99.56	0.7671	89.75	87.05	0.5996	93.17	95.07	0.6834
YIQ	97.51	99.40	0.7917	93.60	87.98	0.6591	95.24	95.30	0.7254
YCbCr	97.72	99.73	0.6568	93.29	87.33	0.5240	95.14	95.28	0.5904
XYZ	96.04	98.42	0.5854	91.34	80.94	0.5597	93.31	92.15	0.5730
U-Net									
RGB	95.78	99.46	0.4034	88.89	91.70	0.3995	91.77	96.67	0.4014
Lab	95.29	98.77	0.4413	88.91	87.93	0.4524	91.58	94.88	0.4469
HSV	98.00	99.66	0.3689	82.27	90.89	0.3573	86.34	96.51	0.3631
YIQ	94.76	99.67	0.4196	83.52	95.02	0.3821	88.22	98.00	0.4009
YCbCr	96.14	99.61	0.5461	87.91	94.00	0.5167	91.35	97.60	0.5314
XYZ	94.29	97.88	0.3372	87.53	86.21	0.3945	90.36	93.69	0.3658





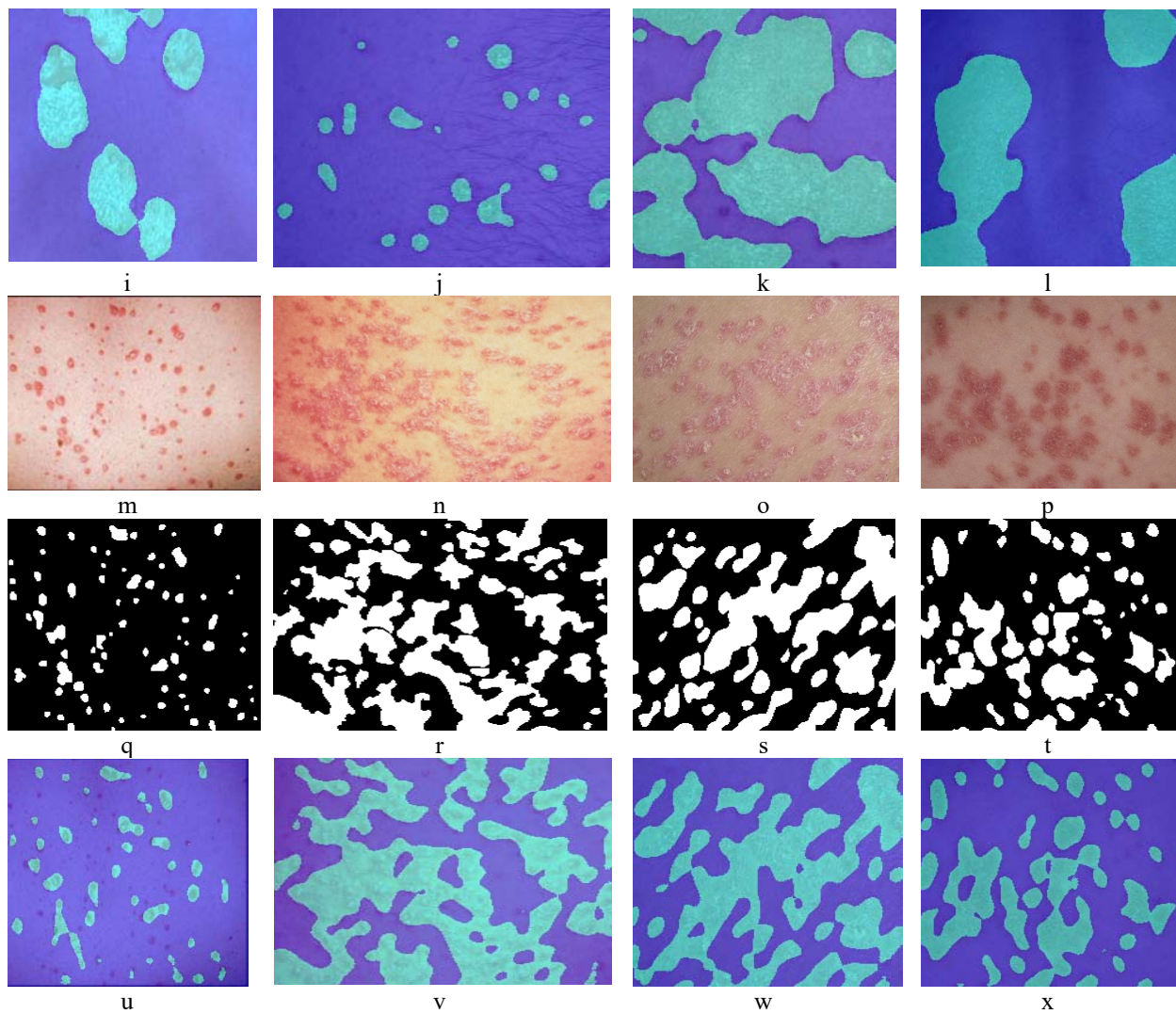


Figure 1 – The segmentation results by proposed CNN:  
 a–d, m–p – the initial psoriasis images; e–h, q–t – the ground-truth images; i–l, u–x – images, segmented by proposed CNN

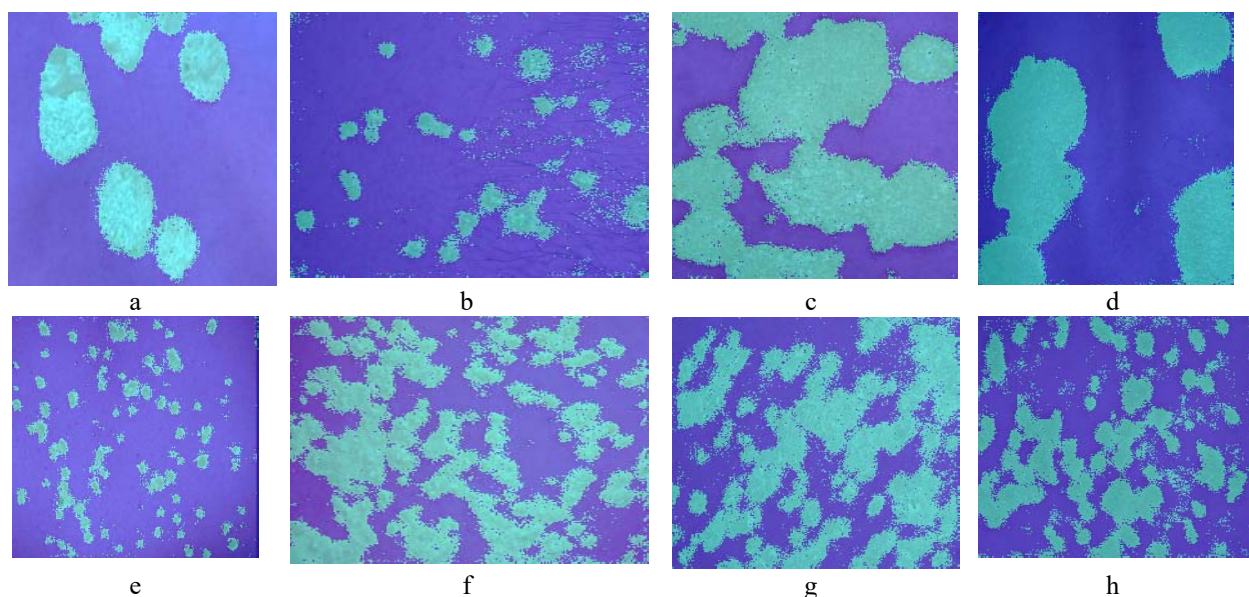


Figure 2 – The segmentation results by U-Net:  
 a–d – the psoriasis images from Figure 1, a–d, segmented by U-Net; e–h – the images from Figure 1, m–p, segmented by U-Net

The accuracy of the psoriasis image segmentation by the proposed CNN, the U-Net, and methods known from the literature are given in Table 3.

Table 3 – The segmentation accuracy of the proposed CNN and methods known from the literature

Reference, publication year, network name	Segmentation accuracy, %
Single Scale Methods	
[7], 2015	79.53%
[8], 2019	87.20%
[19], 2022	92.68%
[6], 2017	95%
Multiscale Methods	
[21], 2017	90%
[9], 2019, PsLSNet	94.8%
[10], 2021, YOLACT	96.6–97.3%
Proposed CNN, RGB images of plaque psoriasis	99.02%
Proposed CNN, RGB images of guttate psoriasis	92.62%
Proposed CNN, Lab images of plaque psoriasis	99.05%
Proposed CNN, Lab images of guttate psoriasis	89.79%
Proposed CNN, YIQ images of plaque psoriasis	98.46%
Proposed CNN, YIQ images of guttate psoriasis	90.79%
[27], 2015, U-Net, RGB images of plaque psoriasis	97.62%
[27], 2015, U-Net, RGB images of guttate psoriasis	90.30%
[27], 2015, U-Net, Lab images of plaque psoriasis	97.03%
[27], 2015, U-Net, Lab images of guttate psoriasis	88.42%
[27], 2015, U-Net, YIQ images of plaque psoriasis	97.22%
[27], 2015, U-Net, YIQ images of guttate psoriasis	89.27%

Fig. 1 illustrates the psoriasis images, the ground-truth images, and the segmentation results, obtained by proposed CNN. At Fig. 2 it is shown the segmentation results, obtained by U-Net which consist of four levels reliaied by 3×3 convolutions. It can be seen from the Fig. 1 that image segmentation by the proposed CNN is characterized by high accuracy in detection of the psoriasis lesion edges.

## 6 DISCUSSIONS

Analysis of the segmentation performance measures given in Table 2 showed the following. It is preferable to segment images of guttate psoriasis lesions in color spaces in which either the red component is isolated separately (RGB), or there is a difference component of red and another color (Lab, YCbCr, YIQ). This is due to the fact that the lesions of guttate psoriasis differ from normal skin mainly by redness level.

Thus, the highest performance of image segmentation of guttate psoriasis lesions by the proposed CNN was obtained using the RGB color space. The Lab, YIQ, YCbCr spaces showed similar results, but compared to the use of the RGB space, TP decreased to 2%, TN decreased by 4%. The presentation of images of guttate psoriasis lesions in xyz and HSV spaces showed the worst results compared to RGB. Specifically, TP is less by 2–4%, TN is less by 4–11%. For images of lesions of plaque psoriasis in the researched color spaces, the segmentation

performance mainly differed within the statistical error. Further, for guttate psoriasis images more often image pixels from the class labeled “Psoriasis lesion” were incorrectly assigned to the class labeled “Normal skin”. For plaque psoriasis images, on the contrary, more often the pixels from the class labeled “Normal skin” were incorrectly assigned to the class labeled “Psoriasis lesion”. This result is due to the relative size of lesions of plaque and guttate psoriasis compared to the normal skin areas in the researched images. Plaque psoriasis lesions occupied a significant area on the researched images; guttate psoriasis lesions are much smaller in area.

The use of U-Net for psoriasis image segmentation [27] instead of the proposed CNN did not improve the segmentation performance.

The proposed CNN is shown a particularly significant advantage in detection of the psoriasis lesion edges. For U-Net the FOM values is less by 17–53% with median 40% for plaque psoriasis images, and the FOM values is less by 2–47% with median 40% for guttate psoriasis images, as compared with the proposed CNN. For all researched psoriasis images the FOM values of U-Net segmentation is less by 10–48% with median 40%.

Next, the Adam method is used for training the proposed CNN. In this case the lower segmentation performance is obtained as compared to stochastic gradient descent. For guttate psoriasis images, TP is less by 3–7%, and TN is decreased by 5–11%, for plaque psoriasis images TP is less by 3%.

Analysis of the segmentation performance of the proposed CNN and the known segmentation methods [6–10, 19–21] given in Table 3 showed the following. The percentage of correctly segmented pixels (accuracy, %) of the proposed CNN exceeds the known methods by up to 13% for guttate psoriasis images and by 4–20% for plaque psoriasis images.

A comparison of processing time was made for the image segmentation by the proposed CNN and the U-Net [27]. The researched psoriasis images were scaled to a size of 224×224 pixels. Then the processing time of the proposed CNN was 3.7–3.8 seconds on average per image when training the network using the stochastic gradient descent method. For a 4-level UNet network, when trained by the Adam method, the average processing time for one image with a 3 × 3 filter was 1.9–2 seconds. The research was performed using an Intel Core i5-7400 processor, 3 GHz CPU, 16GB memory, Windows 10 operating system, 64 bit. Thus, the proposed CNN requires on average 2 times more time to process one image than the 4-layer U-Net network. However, the number of training epochs of the 4-layer U-Net for this segmentation problem is, on average, 30–40% more than is required for training the proposed CNN.

## CONCLUSIONS

The problem of mathematical support elaboration is solved to automate the image processing in dermatological disease diagnostic systems.

**The scientific novelty** of obtained results is that the deep learning convolutional neural network was elaborated to segment color psoriasis images. The improved CNN differs from those known from the literature in that its architecture did not include pooling layers; only convolutional layers, batch normalization, and an activation function were used. The use of the proposed CNN made it possible to enhance the segmentation performance of plaque psoriasis images and, especially, guttate psoriasis images at comparable processing time. There were no abrupt changes in color and illumination, blurring, as well as the complex background, for example, containing clothes or fragments of the interior on the researched psoriasis images. For such images, in cases of slow changes in image illumination, the proposed CNN showed a high segmentation performance, especially at the edges of the lesions.

**The practical significance** of obtained results is that the software realizing the proposed CNN is developed, as well as experiments to research its image segmentation performance are conducted. The experimental results allow to recommend the proposed CNN for use in practice, as well as to determine effective conditions for the application of the proposed CNN.

**Prospects for further research** are to study the performance of the proposed CNN then abrupt changes in color and illumination, blurring, as well as the complex background areas are present on dermatological images, for example, containing clothes or fragments of the interior. In addition, it is advisable to use the proposed CNN in other problems of color image processing to segment statistical or spectral-statistical texture regions on a uniform or textured background.

#### ACKNOWLEDGEMENTS

The author express their deep gratitude to V. N. Krylov, Doctor of Technical Sciences, Professor of the Department of Applied Mathematics and Information Technologies, National University "Odessa Polytechnic" for valuable and constructive advice and comments while working on this paper.

#### REFERENCES

1. Kotvitska A. A., Carlo V. V. Research on the indicators of the prevalence of psoriasis in the world and Ukraine, *Zaporozhye Medical Journal*, 2013, № 3(78), pp. 38–42.
2. Jarad T. S., Dawood A. J. A quantitative technique for systematic monitoring of the treatment efficiency psoriasis lesion, *London Journal of Research in Computer Science and Technology*, 2019, Vol. 19, № 1, pp. 37–46.
3. Velasco J., Pascion C., Alberio J. W. et al. A smartphone-based skin disease classification using MobileNet CNN, *International Journal of Advanced Trends in Computer Science and Engineering*, 2019, Vol. 8, № 5, pp. 2632–2637. DOI: 10.30534/ijatcse/2019/116852019
4. Pal A., Chaturvedi A., Garain U. et. al. Severity grading of psoriatic plaques using deep CNN based multi-task learning, *Pattern Recognition: 23rd International Conference, ICPR, Cancun, Mexico, 4–8 December, 2016 : proceedings*.

- Piscataway, NJ, IEEE, 2016, pp. 1478–1483. DOI: 10.1109/ICPR.2016.7899846
5. Pal A., Garain U., Chandra A., Chatterjee R., Senapati S. Psoriasis skin biopsy image segmentation using deep convolutional neural network, *Computer Methods and Programs in Biomedicine*, 2018, Vol. 59, № 6 (1), pp. 59–69. DOI:10.1016/j.cmpb.2018.01.027
6. Jarad T. S., Dawood A. J. Accurate segmentation of psoriasis diseases images using k-means algorithm based on CIELAB (L\*A\*B) color space, *Journal of Theoretical and Applied Information Technology*, 2017, Vol. 95, № 17, pp. 4201–4210.
7. Pal A., Roy A., Sen K. et al. Mixture model based color clustering for psoriatic plaque segmentation, *Pattern Recognition : Third IAPR Asian Conference, ACPR, Kuala Lumpur, Malaysia, 3–6 November 2015 : proceedings*. Red Hook, NY, IEEE, 2015, pp. 376–380. DOI:10.1109/ACPR.2015.7486529
8. Volkova N. P. Detector quasi-periodic texture segmentation method for dermatologica images processing, *Herald of Advanced Information Technology*, 2019, Vol. 2, № 4, pp. 259–267. DOI: 10.15276/hait.04.2019.2.
9. Dash M., Londhe N. D., Ghosh S., Semwal A., Sonawane R. S. PsLSNet: automated psoriasis skin lesion segmentation using modified U-Net-based fully convolutional network, *Biomedical Signal Processing and Control*, 2019, Vol. 52, pp. 226–237. DOI: 10.1016/j.bspc.2019.04.002
10. Lin G.-S., Lai K.-T., Syu J.-M. et al. Instance segmentation based on deep convolutional neural networks and transfer learning for unconstrained psoriasis skin images, *Applied Sciences*, 2021, Vol. 11, 3155. DOI: 10.3390/app11073155
11. Subbotin S. A. The fractal dimension based quality metrics of data samples and dependence models, *Radio Electronics, Computer Science, Control*, 2017, № 2, pp. 70–81. DOI: 10.15588/1607-3274-2017-2-8
12. Leoshchenko S. D., Oliynyk A. O., Subbotin S. O., Hoffman E. O., Kornienko O. V. Method of structural adjustment of neural network models to ensure interpretability, *Radio electronics, computer science, management*, 2021, № 3, pp. 86–96.
13. Roslan R., Razly I. N. M., Sabri N., Ibrahim Z. Evaluation of psoriasis skin disease classification using convolutional neural network, *International Journal of Artificial Intelligence*, 2020, Vol. 9, № 2, pp. 349–355. DOI: 10.11591/ijai.v9.i2
14. Raza M. A., Liaqat M. S., Shoaib M. A fuzzy expert system design for diagnosis of skin diseases, *Advancements in Computational Sciences: Second International Conference, ICACS, Lahore, Pakistan, 18–20 February, 2019 : proceedings*, IEEE, 2019. pp. 1–7. DOI: 10.23919/ICACS.2019.8689140
15. George Y. M., Aldeen M., Garnavi R. Automatic scale severity assessment method in psoriasis skin images using local descriptors, *Journal of Biomedical and Health Informatics*, 2020, Vol. 24, № 2, pp. 577–585. DOI: 10.1109/JBHI.2019.2910883.
16. Mullangi P., Yarravarapu S. R., Kotipalli P., Urooj S. (ed.). Texture and clustering-based skin disease classification, *Sensors and Image Processing*. Singapore, Springer, 2018, pp. 103–110. DOI: 10.1007/978-981-10-6614-6\_11
17. Munia T. T., Haque I. R., Aymond A. et. al. Automatic clustering-based segmentation and plaque localization in psoriasis digital images, *Healthcare Innovations and Point-of-Care Technologies, Special Topics Conference, IEEE-*

- NIH*, Bethesda Campus, MD, USA, 6–8 November, 2017 : proceedings, IEEE, 2017, pp. 113–116. DOI: 10.1109/HIC.2017.8227597
18. Roy K., Chaudhuri S. S., Ghosh S. et al. Skin disease detection based on different segmentation techniques, *Opto-Electronics and Applied Optics: International Conference, Optronix, Kolkata*. India, 18–20 March, 2019, proceedings, IEEE, 2019, pp. 70–74. DOI: 10.1109/OPTRONIX.2019.8862403
19. Liu X., Krylov V., Jun S. et al. Segmentation and identification of spectral and statistical textures for computer medical diagnostics in dermatology, *Mathematical Biosciences and Engineering*, 2022, Vol. 19, № 7, pp. 6923–6939. DOI: 10.1109/JBHI.2019.2910883.
20. Krylov V. N., Volkova N. P. Vector-difference texture segmentation method in technical and medical express diagnostic systems, *Herald of Advanced Information Technology*, 2020, Vol. 3, № 4, pp. 226–239. DOI: 10.15276/hait.04.2020.2
21. George Y. M., Aldeen M., Garnavi R. Automatic psoriasis lesion segmentation in two-dimensional skin images using multiscale superpixel clustering, *Journal of Medical Imaging*, 2017, Vol. 4, № 4, pp. 59–69. DOI: 10.1117/1.JMI.4.4.044004
22. Psoriasis: Symptoms, Treatment, Images and More – DermNet [Electronic resource]. Access mode: <http://dermnetnz.org/>
23. Polyakova M., Ishchenko A., Volkova N., Pavlov O. The combining segmentation method of the scanned documents images with sequential division of the photo, graphics, and the text areas, *Eastern-European Journal of Enterprise Technologies*, 2018, Vol. 5, № 2, pp. 6–16. DOI: 10.15587/1729-4061.2018.142735
24. Polyakova M. V., Nesteryuk A. G. Improvement of the color text image binarization method using the minimum-distance classifier, *Applied Aspects of Information Technology*, 2021. Vol. 4, № 1, pp. 57–70. DOI: 10.15276/aait.01.2021.5
25. Akinlar C., Topal C. ColorED: color edge and segment detection by edge drawing, *Journal of Visual Communication and Image Representation*, 2017, Vol. 44, pp. 82–94. DOI: 10.1016/j.jvcir.2017.01.024
26. Gonzalez R. C., Woods R. E. Digital Image Processing (4rd Edition). NY, Pearson, 2017, 1192 p.
27. Ronneberger O., Fischer P., Brox T. U-Net: convolutional networks for biomedical image segmentation, *Medical Image Computing and Computer-Assisted Intervention*, 2015, Vol. 9351, pp. 234–241. DOI: 10.48550/arXiv.1505.04597

Received 07.11.2022.

Accepted 01.12.2022.

УДК 004.93

## СЕГМЕНТАЦІЯ ЗОБРАЖЕНЬ ЗГОРТКОВОЮ НЕЙРОННОЮ МЕРЕЖЕЮ БЕЗ ПУЛІНГОВИХ ШАРІВ В СИСТЕМАХ ДІАГНОСТИКИ ДЕРМАТОЛОГІЧНИХ ЗАХВОРЮВАНЬ

**Полякова М. В.** – д-р техн. наук, доцент, професор кафедри прикладної математики та інформаційних технологій Національного університету «Одеська політехніка», Одеса, Україна.

### АНОТАЦІЯ

**Актуальність.** Розглянуто задачу автоматизації процесу сегментації спектрально-статистичних текстурних зображень. Об'єктом дослідження є обробка зображень у системах діагностики дерматологічних захворювань. Метою дослідження є покращення якості сегментації кольорових псоріазних зображень шляхом розробки згорткової нейронної мережі глибокого навчання без пулінгових шарів.

**Метод.** Запропоновано згорткову нейронну мережу для обробки трьохканального псоріазного зображення заданого розміру. Початкові кольорові зображення було масштабовано до заданого розміру, а потім подано на вхідний шар нейронної мережі. Архітектура запропонованої нейронної мережі складається з чотирьох згорткових шарів з пакетною нормалізацією та функцією активації ReLU. Карти ознак із виходу цих шарів передавалися до згорткового шару 1×1 з функцією активації Softmax. Отримані карти ознак подавалися до шару класифікації пікселів зображення. При сегментуванні зображень згорткові та пулінгові шари оцінюють ознаки фрагментів зображення, а повністю зв'язані шари класифікують отримані вектори ознак, виконуючи розбиття зображення на однорідні сегменти. Ознаки сегментації оцінювалися в результаті навчання мережі а допомогою зображень, сегментованих експертом. Отримані ознаки стійкі до завад та спотворень зображень. Об'єднання результатів сегментації в різних масштабах визначається архітектурою мережі. Пулінгові шари включалися в архітектуру запропонованої згорткової нейронної мережі, оскільки вони зменшують розмір карт ознак порівняно з розміром початкового зображення та можуть знизити якість сегментації невеликих псоріазних плям та псоріазних плям складної форми.

**Результати.** Запропоновану згорткову нейронну мережу реалізовано програмно і досліджено при вирішенні задачі сегментації псоріазних зображень.

**Висновки.** Використання запропонованої згорткової нейронної мережі дозволило підвищити якість сегментації зображень пляшкового та крапельного псоріазу, особливо на границях плям. Перспективи подальших досліджень можуть полягати у дослідженні якості сегментації зображень запропонованою згортковою нейронною мережею, якщо на дерматологічних зображеннях присутні різкі зміни кольору та освітленості, розмиття, а також фрагменти складного фону, наприклад, що містять одяг або фрагменти інтер'єру. Доцільно. використовувати запропоновану згорткову нейронну мережу в інших задачах обробки кольорових зображень для сегментації статистичних або спектрально-статистичних текстурних областей на однорідному або текстурованому фоні.

**КЛЮЧОВІ СЛОВА:** псоріазне зображення, сегментація зображення, згорткова нейронна мережа, пулінговий шар, кольоровий простір, глибоке навчання.

## ЛІТЕРАТУРА

1. Kotvitska A. A. Research on the indicators of the prevalence of psoriasis in the world and Ukraine / A. A. Kotvitska, V. V. Carlo // *Zaporozhye Medical Journal*. – 2013. – № 3(78). – P. 38–42.
2. Jarad T. S. A quantitative technique for systematic monitoring of the treatment efficiency psoriasis lesion / T. S. Jarad, A. J. Dawood // *London Journal of Research in Computer Science and Technology*. – 2019. – Vol. 19, № 1. – P. 37–46.
3. A smartphone-based skin disease classification using MobileNet CNN / [J. Velasco, C. Pascion, J. W. Alberio et al.] // *International Journal of Advanced Trends in Computer Science and Engineering*. – 2019. – Vol. 8, № 5. – P. 2632–2637. DOI: 10.30534/ijatse/2019/116852019
4. Severity grading of psoriatic plaques using deep CNN based multi-task learning / [A. Pal, A. Chaturvedi, U. Garain et al.] // *Pattern Recognition: 23rd International Conference, ICPR, Cancun, Mexico, 4–8 December, 2016 : proceedings*. – Piscataway, NJ, IEEE: 2016. – P. 1478–1483. DOI: 10.1109/ICPR.2016.7899846
5. Psoriasis skin biopsy image segmentation using deep convolutional neural network / [A. Pal, U. Garain, A. Chandra et al.] // *Computer Methods and Programs in Biomedicine*. – 2018. – Vol. 59, № 6 (1). – P. 59–69. DOI:10.1016/j.cmpb.2018.01.027
6. Jarad T. S., Dawood A. J. Accurate segmentation of psoriasis diseases images using k-means algorithm based on CIELAB (L\*A\*B) color space / T. S. Jarad, A. J. Dawood // *Journal of Theoretical and Applied Information Technology*. – 2017. – Vol. 95, № 17. – P. 4201–4210.
7. Mixture model based color clustering for psoriatic plaque segmentation / [A. Pal, A. Roy, K. Sen et al.] // *Pattern Recognition : Third IAPR Asian Conference, ACPR, Kuala Lumpur, Malaysia, 3–6 November 2015 : proceedings*. – Red Hook, NY, IEEE: 2015. – P. 376–380. DOI:10.1109/ACPR.2015.7486529
8. Volkova N. P. Detector quasi-periodic texture segmentation method for dermatologica images processing / N. P. Volkova // *Herald of Advanced Information Technology*. – 2019. – Vol. 2, № 4. – P. 259–267. DOI: 10.15276/hait.04.2019.2.
9. PsLSNet: automated psoriasis skin lesion segmentation using modified U-Net-based fully convolutional network / [M. Dash, N. D. Londhe, S. Ghosh et al.] // *Biomedical Signal Processing and Control*. – 2019. – Vol. 52. – P. 226–237. DOI: 10.1016/j.bspc.2019.04.002
10. Instance segmentation based on deep convolutional neural networks and transfer learning for unconstrained psoriasis skin images / [G.-S. Lin, K.-T. Lai, J.-M. Syu et al.] // *Applied Sciences*. – 2021. – Vol. 11, 3155. DOI: 10.3390/app11073155
11. Subbotin S. A. The fractal dimension based quality metrics of data samples and dependence models / S. A. Subbotin // *Radio Electronics, Computer Science, Control*. – 2017. – № 2. – P. 70–81. DOI: 10.15588/1607-3274-2017-2-8
12. Метод структурного донаштування нейромережових моделей для забезпечення інтерпретабельності / [С. Д. Леошенко, А. О. Олійник, С. О. Субботін et al.] // *Радіоелектроніка, інформатика, управління*. – 2021. – № 3. – С. 86–96.
13. Evaluation of psoriasis skin disease classification using convolutional neural network / [R. Roslan, I. N. M. Razly, N. Sabri, Z. Ibrahim] // *International Journal of Artificial Intelligence*. – 2020. – Vol. 9, № 2. – P. 349–355. DOI: 10.11591/ijai.v9.i2
14. Raza M. A. A fuzzy expert system design for diagnosis of skin diseases / M. A. Raza, M. S. Liaqat, M. Shoaib // *Advancements in Computational Sciences: Second International Conference, ICACS, Lahore, Pakistan, 18–20 February, 2019 : proceedings*. – IEEE: 2019. P. 1–7. DOI: 10.23919/ICACS.2019.8689140
15. George Y. M. Automatic scale severity assessment method in psoriasis skin images using local descriptors / Y. M. George, M. Aldeen, R. Garnavi // *Journal of Biomedical and Health Informatics*. – 2020. – Vol. 24, № 2. – P. 577–585. DOI: 10.1109/JBHI.2019.2910883.
16. Mullangi P. Texture and clustering-based skin disease classification / P. Mullangi, S. R. Yarravarapu, P. Kotipalli // *Sensors and Image Processing / S. Urooj (ed.)*. – Singapore, Springer: 2018. – P. 103–110. DOI: 10.1007/978-981-10-6614-6\_11
17. Automatic clustering-based segmentation and plaque localization in psoriasis digital images / [T. T. Munia, I. R. Haque, A. Aymond et al.] // *Healthcare Innovations and Point-of-Care Technologies: Special Topics Conference, IEEE-NIH, Bethesda Campus, MD, USA, 6–8 November, 2017 : proceedings*. – IEEE: 2017. – P. 113–116. DOI: 10.1109/HIC.2017.8227597
18. Skin disease detection based on different segmentation techniques / [K. Roy, S. S. Chaudhuri, S. Ghosh et al.] // *Opto-Electronics and Applied Optics: International Conference, Optronix, Kolkata, India, 18–20 March, 2019 : proceedings*. – IEEE: 2019. P. 70–74. DOI: 10.1109/OPTRONIX.2019.8862403
19. Segmentation and identification of spectral and statistical textures for computer medical diagnostics in dermatology / [X. Liu, V. Krylov, S. Jun et al.] // *Mathematical Biosciences and Engineering*. – 2022. – Vol. 19, № 7. – P. 6923–6939. DOI: 10.1109/JBHI.2019.2910883.
20. Krylov V. N. Vector-difference texture segmentation method in technical and medical express diagnostic systems / V. N. Krylov, N. P. Volkova // *Herald of Advanced Information Technology*. – 2020. – Vol. 3, № 4. – P. 226–239. DOI: 10.15276/hait.04.2020.2
21. George Y. M. Automatic psoriasis lesion segmentation in two-dimensional skin images using multiscale superpixel clustering / Y. M. George, M. Aldeen, R. Garnavi // *Journal of Medical Imaging*. – 2017. – Vol. 4, № 4. – P. 59–69. DOI: 10.1117/1.JMI.4.4.044004
22. Psoriasis: Symptoms, Treatment, Images and More – DermNet [Electronic resource]. – Access mode: <http://dermnetnz.org/>
23. The combining segmentation method of the scanned documents images with sequential division of the photo, graphics, and the text areas / [M. Polyakova, A. Ishchenko, N. Volkova, O. Pavlov] // *Eastern-European Journal of Enterprise Technologies*. – 2018. – Vol. 5, № 2. – P. 6–16. DOI: 10.15587/1729-4061.2018.142735
24. Polyakova M. V. Improvement of the color text image binarization method using the minimum-distance classifier / M. V. Polyakova, A. G. Nesteryuk // *Applied Aspects of Information Technology*. – 2021. – Vol. 4, № 1. – P. 57–70. DOI: 10.15276/aait.01.2021.5
25. Akinlar C. ColorED: color edge and segment detection by edge drawing / C. Akinlar, C. Topal // *Journal of Visual Communication and Image Representation*. – 2017. – Vol. 44. – P. 82–94. DOI: 10.1016/j.jvcir.2017.01.024
26. Gonzalez R. C., Woods R. E. *Digital Image Processing (4rd Edition)* / R. C. Gonzalez, R. E. Woods. – NY : Pearson, 2017. – 1192 p.
27. Ronneberger O. U-Net: convolutional networks for biomedical image segmentation / O. Ronneberger, P. Fischer, T. Brox // *Medical Image Computing and Computer-Assisted Intervention*. – 2015. – Vol. 9351. – P. 234–241. DOI: 10.48550/arXiv.1505.04597

## MACHINE LEARNING DECISION SUPPORT SYSTEMS FOR ADAPTATION OF EDUCATIONAL CONTENT TO THE LABOR MARKET REQUIREMENTS

**Shelehov I. V.** – Associate Professor, PhD, Head of the Department of Computer Science, Sumy State University, Sumy, Ukraine.

**Prylepa D. V.** – Assistant Professor, Department of Computer Science, Sumy State University, Sumy, Ukraine.

**Khibovska Yu. O.** – Postgraduate student of the Department of Computer Science, Sumy State University, Sumy, Ukraine.

**Otroshchenko M. S.** – Postgraduate student of the Department of Computer Science, Sumy State University, Sumy, Ukraine.

### ABSTRACT

**Topicality.** The urgent task of increasing the functional efficiency of machine learning of decision support system (DSS) for assessing compliance with content modern requirements of the educational disciplines of the graduation department based on the results of the employer survey has been solved.

**Goal.** Increasing the functional efficiency of machine learning of DSS for assessing compliance with modern requirements of the educational disciplines content of the first (bachelor's) level specialty educational and professional program based on machine learning and pattern recognition.

**Method.** The method of machine learning of DSS is proposed for adapting the educational content of the graduation department to the labor market requirements. The idea of the method is to maximize the information capacity of the DSS in the machine learning process, which allows in the monitoring mode to guarantee a high full probability of making the correct classification decisions. The method was developed as part of a functional approach to modeling cognitive processes of natural intelligence, which makes it possible to provide DSS with flexibility when retraining the system due to increasing the power of the recognition classes alphabet. The method is based on the principle of maximizing the amount of information in the machine learning process. The modified Kullback information measure, which is a functional of the accuracy characteristics of classification solutions, is considered as a criterion for optimizing machine learning parameters. According to the proposed functional category model, an information-extreme machine learning algorithm was developed based on the hierarchical data structure in the form of a binary decursive tree. The use of such a data structure allows you to automatically divide a large number of recognition classes into pairs of nearest neighbors, for which optimization of machine learning parameters is carried out according to a linear algorithm of the required depth. The geometric parameters of hyperspherical containers of recognition classes were considered as optimization parameters, which were restored in the radial basis of the binary space of Hamming features in the machine learning process. At the same time, the input training matrix was transformed into a working binary training matrix, which was changed in the machine learning process through admissible transformations in order to adapt the input information description of the DSS to the maximum reliability of classification decisions.

**Results.** The informational, algorithmic, and software of the DSS was developed to assess the educational content quality based on the machine analysis results of respondents' answers. Within the framework of the geometric approach, based on the information-extreme machine learning results, highly reliable decisive rules, practically invariant to the multidimensionality of the recognition features space, were constructed based on the hierarchical data structure in the form of a binary decursive tree. The influence of machine learning parameters on the functional effectiveness of machine learning of the DSS was studied on the evaluation example of the educational content of the educational and professional bachelor's program of the specialty 122 Computer Science.

**Conclusions.** The computer modeling results confirm the high functional efficiency of the proposed method of information-extreme hierarchical machine learning and can be recommended for practical use in institutions of higher education to assess compliance with modern requirements of the educational content of graduation departments.

**KEYWORDS:** information-extreme machine learning, functional categorical model, information criterion, hierarchical data structure, decursive tree, educational content.

### ABBREVIATIONS

IEIT is an information-extreme intellectual technology;

DSS is decision support system;

CNN is convolutional neural network.

### NOMENCLATURE

$M$  is a set of recognition classes;

$m$  is a number of the recognition class;

$N$  is a set of recognition features in the structured vector;

$i$  is a number of the recognition feature;

$J$  is a set of structured vectors of recognition features;

$J$  is a number of the structured vector;

$H$  is a set of tiers of decursive tree;

$h$  is a number of the tier of decursive tree;

$S$  is a set of strata of decursive tree;

$s$  is a number of the stratum of decursive tree;

$m_s$  is a serial number of the recognition class in the  $s$ -th stratum;

$\delta_{h,s}$  is a parameter that is equal to half of the control tolerances field of the recognition feature.

$x_{h,s,c}$  is the averaged feature vector of the recognition class neighboring in the stratum

$\delta_H$  is the field of normalized tolerances, which specifies the values range of control tolerances

$E_{h,s,m_s}(d)$  is the informational criterion for optimization of machine learning parameters of DSS to recognize feature vectors of recognition class  $X_{h,s,m_s}^o$ ;

$d$  is a distance measure that is equal to the value of the the recognition class container radius;

$G_E$  is the working area of defining the information criterion;

$G_d$  is the permissible area for changing the radii of the recognition class containers;

$P$  is a set of thematic modules of the educational content, which are evaluated by the respondents;

$T$  is a set of reading information time moments;

$\Omega$  is the recognition features space;

$A$  is the recognition classes alphabet;

$Y$  is the input Euclidean training matrix;

$Y^{|S|}$  is the set of input Euclidean training matrices of recognition classes for all levels of the decursive tree;

$X^{|S|}$  is the set of working binary training matrices for all strata of the decursive tree;

$f_1$  is the training matrix formation operator  $Y$ ;

$f_2$  is the operator for constructing a decursive binary tree  $H$ ;

$f_3$  is the training matrix formation operator  $Y^{|S|}$ ;

$f_4$  is the training matrix formation operator  $X^{|S|}$ ;

$\{k\}$  is a set of machine learning steps;

$\tilde{\mathfrak{R}}^{|M|}$  is a fuzzy division of the feature space into  $M$  recognition classes;

$Y^{|S|}$  is an input training matrix of recognition classes  $S$  strata of decursive tree;

$X^{|S|}$  is a binary training matrix of recognition classes  $S$  strata of decursive tree;

$E$  is a term set of values of the information criterion;

$R$  is an operator of construction of division  $\tilde{\mathfrak{R}}^{|2|}$  of space of signs on recognition classes;

$\psi$  is an operator for testing the basic statistical hypothesis about the affiliation of the vector  $x_{h,s,m_s}$  of the recognition class  $X_{h,s,m_s}^o$ ;

$\gamma$  is an operator of the set formation of exact characteristics for the set system of decisions estimations;

$\varphi$  is an operator for calculating the information criterion for optimizing the machine learning parameters;

$U$  is an operator that regulates the machine learning process;

$K_{h,s,m_s}^{(1)}(d)$  is the number of events in which "own" feature vectors do not belong to recognition class to the container  $X_{h,s,m_s}^o$ ;

$K_{h,s,m_s}^{(2)}(d)$  is the number of events in which "foreign" feature vectors belong to the recognition class container  $X_{h,s,m_s}^o$ ;

$10^{-p}$  is a sufficiently small number, which is entered to avoid division by zero (in practice,  $p = 2$  was accepted);

$x^{(j)}$  is a structured feature vector that can be recognized;

$\mu_{m_s}$  is the function of belongingness of feature vector  $x^{(j)}$  to the recognition class container  $X_{h,s,m_s}^o$ ;

$d_{h,s,m_s}^*$  is the optimal radius of the recognition class container  $X_{h,s,m_s}^o$ .

## INTRODUCTION

Establishing a stable connection between the graduation department of a higher education institution and employers is a necessary condition for improving the quality of the educational process. The main way to establish such a connection is to survey employers and specialists in the relevant field of knowledge. The development of modern information and communication technologies makes it possible to automate data collection. At the same time, in practice, the results of the respondents survey are processed using the multidimensional statistical analysis methods. Such methods to ensure statistical stability and homogeneity require large volumes of input data, which requires organizers to spend significant time and money on processing and analyzing survey results.

**The object of research** is the process of sustainable monitoring of the education quality by creation of the decision support system (DSS) for adapting the educational content of the graduation department to the labor market requirements on the machine learning and pattern recognition basis.

**The subject of research** is the method of hierarchical information-extreme machine learning of DSS for adapting the educational content of the graduation department to the labor market requirements.

**The purpose of the work is** to increase the functional efficiency of the DSS machine learning during automatically forming an input training matrix based on the respondents survey results, building decisive rules according to the optimal (here and in the text in the text in the informational sense) machine learning parameters, and in the monitoring mode assessing the compliance of the educational content of the graduation department with modern requirements.

## 1 LITERATURE REVIEW

The article considers the method of hierarchical information-extreme machine learning of DSS for adapting the educational content of the graduation department to the labor market requirements. According to European educational standards, the quality of education is determined by the benefits that both employers and graduates of a higher education institution will receive [1, 2]. One of the main ways of organizing sustainable monitoring of the education quality is the creation of the DSS for adapting the educational content of the graduation department to the labor market requirements [3]. A traditional approach to processing and analyzing the stakeholders survey results and specialists in the knowledge relevant fields is the application of decision-making systems existing in social communications using multidimensional statistical analysis methods [4–6]. At the same time, the disadvantages of statistical decision-making methods are the need for large volumes of data, ensuring their statistical stability and homogeneity. A promising direction for increasing the functional efficiency of computerized systems for assessing the quality of education is the use of intelligent information technologies for data analysis [7–9]. Works [10, 11] give examples of creating expert systems for evaluating the quality of the educational process based on fuzzy logic. The main disadvantages of such systems are that they are inflexible and do not provide feedback in the monitoring mode between the graduation department, employers and students of higher education of various forms of education. The further development of computerized systems for assessing the education quality is the development of the scientific and methodological foundations of the information synthesis of DSS, capable of automatically forming an input information description and identifying regularities based on machine learning and pattern recognition [12]. CNN [13, 14] is the most common among known intelligent information technologies for data analysis, but the main drawback of CNN is its sensitivity to the multidimensionality of the feature space and the recognition classes alphabet. Works [15, 16] consider the use of extractors based on artificial neural networks to reduce the impact of input data multidimensionality, but this approach is associated with the information loss possibility.

Works [17–19] consider the use of fuzzy neural networks for functional diagnosis, but at the same time there is also the multidimensionality problem, which significantly limits the capabilities of the fuzzy logic apparatus.

The main scientific and methodological reasons that complicate the use of CNN for the information synthesis of the DSS for the education quality assessment are:

- arbitrary initial conditions of the evaluation process;
- intersection in the recognition classes features space that characterize the corresponding levels of educational content quality;
- multidimensionality of the signs dictionary;

– the impact on the machine assessment of the education quality of uncontrollable disturbing factors, for example, man-made disasters, an unfavorable epidemiological situation or the introduction of martial law.

A promising way to reduce the multidimensionality impact of the recognition features dictionary is the use of machine learning methods, based on the results of which decisive rules are built within the framework of a geometric approach [20]. Among such methods, the information technologies of the support vector method [21, 22] deserve attention, but the functional effectiveness of the algorithms of this method significantly depends on the degree of recognition classes intersection in the feature space. This shortcoming is not present in machine learning methods, which are developed within the framework of the so-called information-extreme intelligent data analysis technology (IEIT) created at Sumy State University (Ukraine) [23–25], the methods of intelligent data analysis proposed within this technology are based on the principle maximizing the information capacity of the system in the process of its machine learning. The work [25] considered information-extreme machine learning based on a hierarchical data structure. But in the case of hierarchical machine learning DSS to assess the educational content quality, it is necessary to take into account such a feature as the presence of an ordered alphabet of recognition classes that characterize different levels of educational content quality when building a hierarchical data structure in the form of a decursive tree.

The purpose of the article is to develop an algorithm of hierarchical information-extreme machine learning of DSS and to verify it on the example of assessing compliance with modern requirements of the educational content of the educational and professional bachelor's degree program of the "Computer Science" specialty.

## 2 PROBLEM STATEMENT

Let's consider within the framework of IEIT the formalized setting of the information synthesis task of DSS to assess of the educational content quality and professional program of the specialty, which is used to train students of higher education at the graduation department.

It is necessary to build a hierarchical data structure in the form of a decursive binary tree  $\{X_{h,s,m}^o \mid h = \overline{1, H}; s = \overline{1, S_h}; m_s = \overline{1, 2}\}$  for the given alphabet  $\{X_m^o \mid m = \overline{1, M}\}$  of recognition classes that characterize the levels of educational content quality assessment according to the appropriate system.

According to the concept of IEIT, transform the input training matrices of the recognition classes of each layer into the corresponding working binary matrices specified in the Hamming space. At the same time, let the DSS machine learning parameters be set, which, for example,



for recognition class  $X_{h,s,m_s}^o$  are represented in the form of a structured vector

$$g_{h,s} = \langle x_{h,s,m_s}^o, d_{h,s,m_s}, \delta_{h,s} \rangle. \quad (1)$$

The number of optimization parameters in expression (1) sets the second depth level of information-extreme machine learning, since vector  $x_{h,s,m_s}$  depends on parameter  $\delta_{h,s}$  depends on parameter.

Machine learning parameters are limited:

- a)  $d_{h,s,m_s} \in [0; d(x_{h,s,m_s} \oplus x_{h,s,c})]$ ;
- b)  $\delta_{h,s} \in [0; \delta_H / 2]$ .

In the machine learning process of DSS, it is necessary to:

1) optimize the parameters of the vector (1) by finding the global maximum of the alphabetically averaged recognition classes of the  $s$ -th stratum  $h$ -th tier of the information criterion:

$$\bar{E}_{h,s} = \frac{1}{2} \sum_{m_s=1}^2 \max_{G_E \cap G_d} E_{h,s,m_s}(d), \quad (2)$$

2) in the information-extreme machine learning process, build highly reliable decision rules based on the optimal (here and in the text in the informational sense) geometric parameters of the recognition classes containers;

3) during the operation of the DSS in the monitoring mode, determine whether the recognized feature vector belongs to the corresponding class from the given alphabet.

### 3 MATERIALS AND METHODS

Information-extreme machine learning of DSS was carried out using a hierarchical data structure in the form of a decursive binary tree [24]. With such a structure, the attribute from the top of the stratum of the higher tier is transferred to one of the vertices of the child stratum  $s$ , the lower tier. At the same time, the training matrix of the corresponding recognition class is considered as an attribute of the vertex.

Construction of a decursive binary tree was carried out according to the scheme:

1) alphabet  $\{X_m^o | m = \overline{1, M}\}$  of ordered recognition classes is divided into two groups, which define two branches of the decursive tree, respectively;

2) as attributes of the vertices of the upper (first according to dendrographic classification) tier of the decursive tree, the training matrices of the boundaries for each of the recognition classes groups are selected;

3) the attributes of the strata of the upper tier are transferred to the vertices of the corresponding strata of the lower tier;

4) the strata of the lower tiers of each branch of the tree contain, in addition to the training matrix transported from the upper tier, also the training matrix of the nearest neighboring recognition class in its group;

5) the construction of the tree continues until the final strata are formed, which contain the training matrices of all recognition classes.

Thus, the binary decursive tree built according to the above scheme divides the given recognition classes alphabet into strata, each of which contains two nearest neighboring classes, which allows applying a linear information-extreme machine learning algorithm for each final stratum. At the same time, the construction of error-free decisive rules based on the training matrix is achieved by optimizing additional parameters of the DSS functioning.

The functional categorical model of the information-extreme machine learning of DSS is considered in the form of a directed graph, the edges of which are the mapping operators of the corresponding sets. At the same time, the input information description of the DSS is given by the structure

$$I = \langle P, T, \Omega, A, Y, H, Y^{[S]}, X^{[S]}; f_1, f_2, f_3, f_4 \rangle.$$

At the same time, the Cartesian product  $P \times T \times \Omega \times Z$  is considered as a source of information.

The functional categorical model of information-extreme machine learning of DSS with optimization of machine learning parameters (1) is shown in Fig. 1.

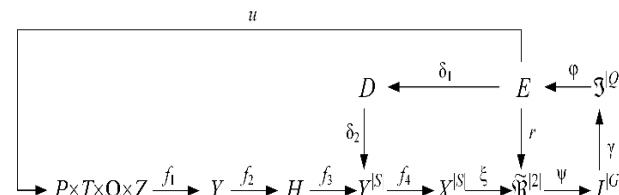


Figure 1 – Functional categorical model of machine learning of DSS

In Figure 1 operator  $\xi$  maps the structured binary feature vectors of training matrices  $X^{[S]}$  to a fuzzy, in the general case, partition of  $\tilde{\mathfrak{R}}_{h,s}^{[2]}$  recognition classes of each stratum of the decursive tree. The classification operator  $\psi: \tilde{\mathfrak{R}}^{[2]} \rightarrow I^{[G]}$ , where  $I^{[G]}$  – is a set of  $G$  statistical hypotheses, checks the basic statistical hypothesis that the vector  $x_{h,s,m_s}^{(j)}$  belongs to the fuzzy recognition class

$X_{h,s,m_s}^o$ . By evaluating statistical hypotheses, operator

$\gamma$  forms a set of accuracy characteristics  $\mathfrak{Z}^{[Q]}$ , where  $Q = G^2$  is the number of accuracy characteristics. Operator  $\varphi$  calculates a set of values of information criterion  $E$ , which is a functional of accuracy

characteristics. The optimization contour of the geometric parameters of partition  $\tilde{\mathfrak{R}}^{|M|}$ . The optimization contour of the geometric parameters of partition  $r$ , which at each step of machine learning restores the recognition classes containers in the radial basis of the Hamming feature space. The optimization contour of parameter  $\delta_{h,s}$  of the control tolerances field includes the term set  $D$ , the elements of which are the allowable values of control tolerances for recognition features. At the same time, operator  $\delta_1$  changes the control tolerances, and operator  $\delta_2$  changes the quantization levels of the vector recognition features of the training matrices  $Y^{|S|}$ . Operator  $u$  regulates the machine learning process.

According to the functional category model (Fig. 1). We will present the algorithm of information-extreme machine learning of DSS in the form of a two-cycle iterative procedure for finding the global maximum of an information criterion in the working area of determining its function:

$$\delta_{h,s}^* = \arg \max_{G_\delta} \{ \max_{G_E \cap G_d} \bar{E}_{h,s}(d) \}. \quad (3)$$

The internal loop of the procedure (3) implements the basic machine learning algorithm, the purpose of which is to calculate at each step of learning the optimization information criterion  $\bar{E}_{h,s}^{(k)}$  and search for the maximum value in the working area of determining its function.

Since the values of recognition features have the same measurement scale, a machine learning algorithm was implemented with parallel optimization of the control tolerances system, according to which control tolerances are changed for all recognition features simultaneously at each step of machine learning.

The input information for the machine learning algorithm is arrays of training matrices  $\{y_{h,s,m_s}^{(j)}\}$  and a system of fields of normalized tolerances  $\{\delta_{H,i}\}$  for recognition features, which sets the values range of the corresponding control tolerances.

Consider the optimization scheme of machine learning parameters for the recognition classes of the  $s$ -th stratum of the  $h$ -th tier of the decursive tree:

1) the recognition class counter of the  $s$ -th stratum of the  $h$ -th tier of the decursive tree is reset:  $m_s := 0$ ;

2)  $m_s := m_s + 1$ ;

3) the parameter change step counter is reset to zero  $\delta_{h,s} : l := 0$ ;

4) the counter  $l := 1 + 1$  is started and the lower  $A_{HK,i}$  and upper  $A_{BK,i}$  control tolerances are calculated for all recognition features:

$$\{A_{HK,i}[l] = y_{h,s,i} - l\}; \{A_{BK,i}[l] = y_{h,s,i} + l\}; i = \overline{1, N},$$

5) the basic algorithm of information-extreme machine learning is implemented, the tasks of which are to calculate at each step of learning the value of the information criterion (2) and search for the maximum value of the criterion in the working area of determining its function;

6) the optimal radius of the recognition class container  $X_{h,s,m_s}^o$  is determined

$$d_{h,s,m_s}^* = \arg \max_{G_E \cap G_d} E_{h,s,m_s}(d),$$

7) if  $l \leq \delta_H / 2$ , then point 4 is fulfilled, otherwise – point 8;

8) if  $m_s \leq 2$ , then point 2 is fulfilled, otherwise point 9;

9) the average value of information criterion  $\bar{E}_{h,s}$  and the optimal parameter of the control tolerances field  $\delta_{h,s}^*$  are calculated;

10) STOP.

Since the machine learning of DSS takes place according to the hierarchical data structure, an algorithm was implemented, the main stages of which are:

1) zeroing of the data structure tiers counter:  $h := 0$ ;

2) initialization of the data structure tier counter:  $h := h + 1$ ;

3) zeroing of the stratum counter:  $s := 0$ ;

4) initialization of the tier counter:  $s := s + 1$ ;

5) for each  $s$ -th stratum of the  $h$ -th tier of the decursive tree, an information-extreme algorithm of machine learning with parallel optimization of control tolerances for recognition features is implemented, which calculates:

a) the maximum value of the information criterion  $\bar{E}_{h,s}$  averaged for the stratum;

b) averaged for the stratum  $\{x_{h,s,m_s}^*\}$  recognition classes of the  $s$ -th stratum of the  $h$ -th tier;

c) optimal radii of  $\{d_{h,s,m_s}^*\}$  recognition classes of the  $s$ -th stratum of the  $h$ -th tier;

d) optimal parameter  $\delta_{h,s}^*$  of the control tolerances field for recognition classes features of the  $s$ -th stratum of the  $h$ -th tier of the decursive tree;

6) if  $s \leq S_h$ , where  $S_h$  is the number of strata on the  $h$ -th tier, then point 4 is fulfilled, otherwise point 7;

7) if  $h \leq h_{\max}$ , where  $h_{\max}$  is the number of the decursive tree levels, then point 2 is fulfilled, otherwise point 8;

8) the maximum value of information criterion  $\bar{E}_H$  averaged over all strata of the decursive tree is calculated;

9) a structured set of decisive rules is formed for all recognition classes;

10) STOP.

As a criterion for optimizing machine learning parameters, the modified Kullback information measure was considered, which for two-alternative solutions and equally likely hypotheses has the form [25]

$$E_{h,s,m_s}(d) = \frac{\left[ n - (K_{h,sw,m_s}^{(1)}(d) + K_{h,s,m_s}^{(2)}(d)) \right]}{n} \times \log_2 \left\{ \frac{2n + 10^{-P} - \left[ (K_{h,sw,m_s}^{(1)}(d) + K_{h,s,m_s}^{(2)}(d)) \right]}{\left[ K_{h,s,m_s}^{(1)}(d) + K_{h,s,m_s}^{(2)}(d) \right] + 10^{-P}} \right\}. \quad (4)$$

Based on the optimal geometric parameters of the recognition classes containers obtained in the machine learning process, decisive rules are constructed, which will be presented in the form

$$(\forall X_{m,h,s}^o \in \mathfrak{R}^{|M|})(\forall x^{(j)} \in \mathfrak{R}^{|M|}) \{ \text{if } [(\mu_{m_s} > 0) \& \& (\mu_m = \max_{\{m\}} \{ \mu_{m_s} \mid m_s = \overline{1, 2} \})] \text{ then } x^{(j)} \in X_{h,s,m_s}^o \text{ else } x^{(j)} \notin X_{h,s,m_s}^o \}. \quad (5)$$

In expression (5), the membership function for the hyperspherical container of recognition class  $X_{h,s,m_s}^o$  is determined by the formula

$$\mu_{m_s} = 1 - \frac{d(x_{h,s,m_s}^* \oplus x^{(j)})}{d_{h,s,m_s}^*}. \quad (6)$$

In expression (6), the code distance between the optimal averaged feature vector  $x_{h,s,m_s}^*$  and the recognized feature vector  $x^{(j)}$ , is denoted as  $d(x_{h,s,m_s}^* \oplus x^{(j)})$ .

Thus, the geometric decisive rules (5) built in the information-extreme machine learning process of DSS are characterized by low computational complexity and are practically invariant to the multidimensionality of the recognition features dictionary.

#### 4 EXPERIMENTS

The implementation of the above information-extreme machine learning algorithm of DSS was carried out on the example of assessing the compliance with the labor market requirements of the educational content of the bachelor's level educational and professional program in the specialty "Computer Science", which is taught to students of Sumy State University. The input training matrix was formed by simulating the evaluations of the educational disciplines thematic modules, which were

© Shelehov I. V., Prylepa D. V., Khibovska Yu. O., Otroshchenko M. S., 2023  
DOI 10.15588/1607-3274-2023-1-6

randomly generated according to the normal distribution of probabilities and displayed on a stobal scale. Since the thematic modules were evaluated according to the European rating scale, the stobal scale was previously divided into six fuzzy intervals, which specified the areas of the recognition classes:

- 1) recognition class  $X_A^o$  included feature vectors that fell within the range from 88 to 100 points;
- 2) for recognition class  $X_B^o$  – from 81 to 91 points;
- 3) for recognition class  $X_C^o$  – from 74 to 84 points;
- 4) for recognition class  $X_D^o$  – from 67 to 77 points;
- 5) for recognition class  $X_E^o$  – from 59 to 69 points;
- 6) for recognition class  $X_F^o$  – from 50 to 62 points.

The simulation training matrix for each recognition class consisted of 40 syllabuses of educational disciplines of the educational and professional program structured by thematic modules. Each vector consisted of 144 features, the number of which was equal to the thematic modules number. In addition, according to the educational and professional program of the specialty, educational disciplines were divided into seven blocks, which included: general scientific, fundamental, design, technological organizational, humanitarian and selective educational disciplines.

For a given recognition classes alphabet, a hierarchical data structure in the form of a decursive tree was built according to the above algorithm (Fig. 2).

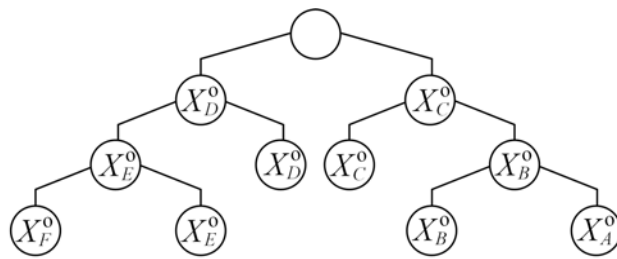


Figure 2 – Hierarchical data structure in the form of a decursive binary tree

According to the decursive tree (Fig. 2), information-extreme machine learning was implemented according to procedure (3). Optimization of machine learning parameters specified by vector (1) was carried out for the two nearest neighboring recognition classes of each stratum. Based on the optimal geometric parameters of the recognition class containers obtained in the machine learning process, decisive rules (5) were constructed. Since each recognition class belongs to two strata, according to the minimum-distance principle of pattern recognition theory, the geometric parameters of the recognition class whose container had the minimum optimal radius were taken as optimal.

According to the machine learning results of DSS, the values of the averaged information criteria (4) were analyzed for each stratum in order to determine the need

to increase the depth of machine learning by optimizing additional parameters of the DSS operation, including the formation parameters of the input information description.

### 5 RESULTS

In the programmatic implementation process of information-extreme machine learning of DSS according to the iterative procedure (3), the optimal parameters of the recognition classes for each of the strata of the decursive binary tree were determined (Fig. 2). As an example, Figure 3 shows the dependency graph of the averaged information criterion (4) for the stratum of the first (upper according to dendrographic classification) tier of the decursive tree, obtained by the machine learning results.

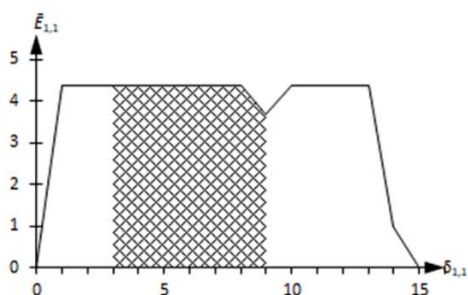


Figure 3 – Dependency graph of the information criterion on the parameter of the control tolerances field for the strata of the first tier

In Fig. 3 double hatching indicates the working area of the definition of the criterion function (4), in which the errors of the first and second kind are less than the first and second reliabilities, respectively. The analysis of Figure 3 shows that the global maximum of the optimization criterion is reached in the area of the plateau-type graph. In this case, the optimal parameter is determined under the condition of the minimum value of the so-called coefficient of recognition classes intersection [23]

$$\eta = \frac{d_{h,s,m_s}^*}{d(x_{h,s,1}^* \oplus x_{h,s,2}^*)} \rightarrow \min. \quad (7)$$

In expression (7), the intercenter code distance between the recognition classes and the  $X_{h,s,1}^o$  and  $X_{h,s,2}^o$   $s$ -th stratum of the  $h$ -th tier of the decursive tree is denoted as  $d(x_{h,s,1}^* \oplus x_{h,s,2}^*)$ .

After checking the fulfillment of condition (7), the optimal values interval of the control tolerance field parameter  $\delta_{1,1}^* = 5$  was determined, from which the value  $\overline{E}_{1,1}^* = 4.50$  was selected. At the same time, the averaged information optimization criterion reaches its maximum

value at parameters  $n=40$  and  $p=2$  specified in formula (4).

Since decisive rules (5) are built within the framework of a geometric approach, they require knowledge of the optimal geometric parameters of recognition class containers obtained in the machine learning process. Figure 4 shows dependency graphs of criterion (4) on the radii of the recognition classes containers of the first tier of the decursive tree (Fig. 2).

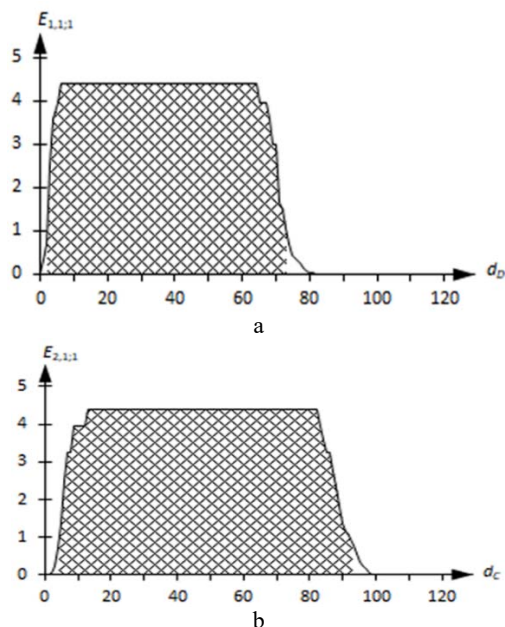


Figure 4 – Dependency graphs of the information criterion (4) on the radii of the recognition classes containers of the first stratum of the second tier: a – class  $X_D^o$ ; b – class  $X_C^o$

The optimal radii of the recognition classes containers of the first stratum according to the condition (7) are equal to any number from the intervals:  $d_D^* \in [38,39]$  (hereafter in code units) for class  $X_D^o$  and  $d_C^* \in [24,27]$  for class  $X_C^o$ . According to the minimum-distance principle of the pattern recognition theory, the minimum radii were chosen:  $d_D^* = 38$  and  $d_C^* = 24$ .

Fig. 5 shows dependency graph of the averaged information criterion (4) on parameter  $\delta$ , obtained during machine learning based on the training matrix of recognition classes of the first stratum of the second tier of the decursive tree (Fig. 2).

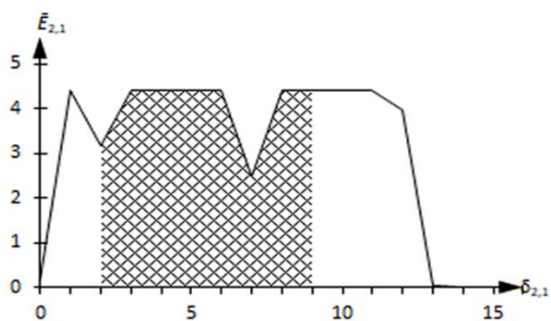


Figure 5 – Dependency graph of the information optimal parameter of machine learning of the criterion on the parameter of the control tolerances field for the first stratum of the second tier

When condition (7) is fulfilled, the optimal parameter of the control tolerances field for recognition classes of the first stratum of the second tier of the decursive tree (Fig. 2) is equal to  $\delta_{2,1}^* = 4$  at the maximum value of the criterion  $\bar{E}_{2,1}^* = 4.50$ .

Figure 6 shows dependency graphs of criterion (4) on the radii of the recognition classes containers of the first stratum of the second tier of the decursive tree (Fig. 2).

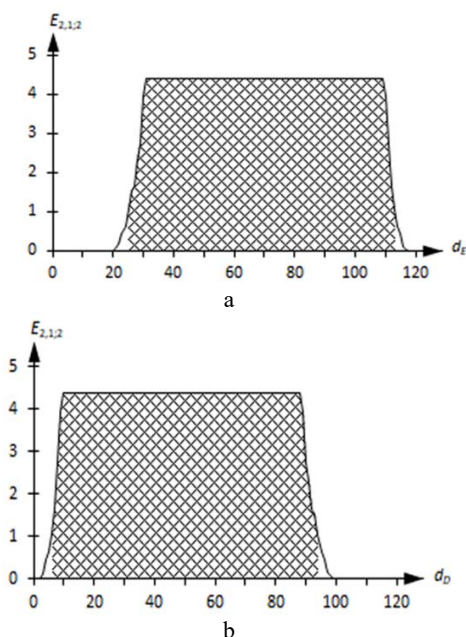


Figure 6 – Dependency graphs of the information criterion (4) on the radii of the recognition classes containers of the first stratum of the second tier: a – class  $X_{2,1,1}^o (X_E^o)$ ; b – class

$$X_{2,1,2}^o (X_C^o)$$

The optimal radii of the recognition classes containers of the first stratum of the second tier according to condition (7) are equal to  $d_{2,1,1}^* = 46$  for class  $X_E^o$  and  $d_{2,1,2}^* = 32$  for class  $X_D^o$ . Similarly, the optimal machine learning parameters for other recognition classes were

determined. Table 1 shows the results of information-extreme machine learning of DSS according to the hierarchical structure (Fig. 2).

Table 1 – Optimal machine learning parameters

Стратум (h,s)	$\delta_{h,s}^*$	$d_{h,s,1}^*$	$d_{h,s,2}^*$
1,1	$\delta_{1,1}^* = 5$	$d_{1,1,1}^* = d_D^* = 38$	$d_{1,1,2}^* = d_C^* = 24$
2,1	$\delta_{2,1}^* = 4$	$d_{h,s,1}^* = d_E^* = 46$	$d_{2,1,2}^* = d_D^* = 32$
2,2	$\delta_{2,2}^* = 5$	$d_{2,2,1}^* = d_C^* = 52$	$d_{2,2,2}^* = d_B^* = 16$
3,1	$\delta_{3,1}^* = 5$	$d_{3,1,1}^* = d_F^* = 64$	$d_{3,2,2}^* = d_E^* = 12$
3,2	$\delta_{3,2}^* = 4$	$d_{3,2,1}^* = d_B^* = 57$	$d_{3,2,2}^* = d_A^* = 16$

Comparing in Table 1 the radii of the recognition classes containers  $X_D^o$  for the strata of the upper and lower tiers, a smaller value of  $d_D^* = 32$ , should be taken as the optimal one, since class  $X_E^o$  is the nearest neighbor for it. Similarly, for recognition class  $X_C^o$  the permissible optimal radius is equal to  $d_C^* = 24$ , for class  $X_B^o - d_B^* = 16$  and for class  $X_E^o - d_E^* = 12$ .

When calculating the membership function (6), vectors  $x_{h,s,m_s}^*$  were determined by optimal control tolerances for recognition features.

## 6 DISCUSSION

The advantages of building a hierarchical data structure in the form of a decursive binary tree are shown on the example of information-extreme machine learning of DSS for adapting the educational content of the educational and professional program of the second (bachelor) level to the labor market requirements. Thanks to the proposed structure, it was possible to automatically divide a set of recognition classes into pairs of nearest neighbors. In addition, the a priori structuring of the recognition classes due to the corresponding knowledge level assessment system is taken into account. As a result, there is no need to form a variational series of recognition classes, which is an advantage of the training required to build a decursive binary tree. The use of a hierarchical data structure in the form of a decursive binary tree, in addition, allows for each stratum to implement information-extreme machine learning according to a linear algorithm with the required level of depth. Unlike neuro-like structures in IEIT methods, the depth level of information-extreme machine learning is determined by the number of optimization parameters that affect the functional efficiency of the intelligent system. This approach corresponds to a greater extent to the mechanism of making classification decisions by natural intelligence. It is also justified to use information criterion (4) in classification tasks as an optimization criterion,

which in the logical and epistemological aspect is considered as a measure of the diversity of recognition classes.

The analysis of the obtained results of the information-extreme machine learning of DSS for adapting the educational content to the labor market requirements shows that for each layer of the decursive tree it was possible to build error-free decisive rules according to the training matrix. At the same time, in accordance with the minimum-distance principle of pattern recognition theory, the radii of the recognition classes containers determined relative to the nearest neighboring class were considered optimal.

The specifics of the development of algorithmic and software of the DSS for adapting educational content to the labor market requirements is the use of simulation modeling for the formation of the input training matrix. Such an approach is due to the large material and time costs of obtaining the results of the respondents survey in the absence of the DSS. However in the future, when the DSS is functioning in the monitoring mode, as archival representative data accumulates, it becomes possible to retrain the system based on real data. Further research will be aimed at expanding the functional capabilities of the DSS in order to evaluate blocks and individual academic disciplines. As a result, graduation departments will be able to quickly adjust the educational content to modern requirements.

### CONCLUSIONS

The actual problem of improving the quality of education is being solved by building an information and communication system for adapting the educational content of the graduation department of the university to the requirements of the labor market.

**The scientific novelty** of the obtained results is that the method of information-extreme machine learning based on the hierarchical data structure in the form of a decursive binary tree is proposed for the first time. The method automatically divides the alphabet of high-power recognition classes into pairs of nearest neighbors. It performs two-class machine learning of DSS, ensuring high reliability with a minimum depth of machine learning. As a result, error-free decision rules based on the training sample were built in the monitoring mode of DSS. It allows the total probability of correct classification decisions to be close to the maximum limit.

**The practical significance** of the obtained results is that the developed DSS software allows for quickly adjusting the educational content of the graduation department of the university, taking into account the requirements of the labor market, with minimal material and time costs.

**Prospects for further research** consist in simplifying the formation of the input information description of the DSS by automatically reading the thematic modules from the syllabuses of the educational disciplines of the specialty of the corresponding level of training of higher education applicants.

### ACKNOWLEDGEMENTS

The work was carried out in the laboratory of intelligent systems of the Department of Computer Sciences of Sumy State University with the funds of the state budget scientific research theme “Information technology of autonomous navigation of unmanned aerial vehicles by terrestrial natural and infrastructural landmarks” (№ DR 0122U000786).

### REFERENCES

1. Bates T. National strategies for e-learning in post-secondary education and training. UNESCO, 2001, 132 p.
2. Sharable Content Object Reference Model (SCORM) Version 1.2 / Advanced Distributed Learning, 2001, 170 p.
3. Agapova M. O. Pro problemu optymizatsii navchalnoho materialu, *Teorija i praktyka upravlinnja social'nymy systemamy*, 2012, № 1, pp. 34–39.
4. Nagy A., Bruck P., Buchholz A., Karssen Z., Zeffass A. eds E-Content: Technologies and Perspectives for the European Market. Berlin, Springer-Verlag, 2005, pp. 79–96.
5. Dovgyj S. O., Bidyuk P. I., Trofymchuk O. M. Decision support systems based on statistical and probabilistic procedures. Kyiv, Logos, 2014, 419 p.
6. Boyd S., Parikh N., Chu E. et al. Distributed optimization and statistical learning via the alternating direction method of multipliers, *Foundations and Trends R in Machine Learning*, 2011, Vol. 3(1), pp. 1–122.
7. Xu G., Zong Y., Yang Y. Z. Applied Data Mining, CRC Press, 2013, 284 p.
8. Dua S., Du X. Data Mining and Machine Learning in Cybersecurity, Auerbach Publications, 2011, 256 p.
9. Sklyar V. V., Kharchenko V. S. Case-Based and Project-Based Methods for Effective E-learning in ICT Safety and Security, *ICTERI Workshops*, 2020, pp. 157–170.
10. Stuart Russell, Peter Norvig Intelligence artificielle, 2e éd. (French translation of Artificial Intelligence: A Modern Approach, second edition). Paris, Pearson Education France, 2006.
11. Sivanandam S. N., Sumathi S., Deepa S. N. Introduction to fuzzy logic using MATLAB. Springer, 2007, 430 p.
12. Watt J., Borhani R., Katsagelos A. Machine Learning Refined: Foundations, Algorithms, and Applications. 2nd Edition, Cambridge University Press, 2020, 594 p.
13. Wu S.-J., Gebraeel N., Lawley M. A. et al. A neural network integrated decision support system for condition-based optimal predictive maintenance policy. *Systems, Man and Cybernetics, Part A: Systems and Humans. IEEE Transactions*, 2007, Vol. 37, № 2, pp. 226–236.
14. Radovic M., Adarkwa O., Wang Q. Object recognition in aerial images using convolutional neural networks, *J. Imaging*, 2017, Vol. 3, № 2. DOI: 10.3390/jimaging3020021.
15. Alpaydin Ethem. Introduction to Machine Learning. London, The MIT Press, 2010, 110 p.
16. Moskalenko V. V., Korobov A. G. Information-extreme algorithm of the system for recognition of objects on the terrain with optimization parameter feature extraction, *Radio Electronics, Computer Science, Control*, 2017, №2, pp. 38–45.
17. Shi Y., Mizumoto M. An improvement of neuro-fuzzy learning algorithm for tuning fuzzy rules, *Fuzzy sets and systems*, 2001, Vol. 118, № 2, pp. 339–350.
18. Subbotin S. A. The neuro-fuzzy network synthesis and simplification on precedents in problems of diagnosis and

- pattern recognition, *Optical Memory and Neural Networks (Information Optics)*, 2013, Vol. 22, № 2, pp. 97–103. DOI: 10.3103/s1060992x13020082.
19. Efendigil T., Önüt S., Kahraman C. A decision support system for demand forecasting with artificial neural networks and neuro-fuzzy models: a comparative analysis, *Expert Systems with Applications*, 2009, Vol. 36, № 3, pp. 6697–6707.
20. Rahimi A., Recht B. Random features for large-scale kernel machines, *Advances in Neural Information Processing Systems*, 2008, pp. 1177–1184.
21. Chapelle O. Training a support vector machine in the primal, *Neural Computation*, 2007, Vol. 19, №5, pp. 1155–1178.
22. Ivanciuc O. Applications of Support Vector Machines in Chemistry, *Reviews in Computational Chemistry*, 2007, Vol. 23, pp. 291–400.
23. Dovbysh A. S., Budnyk M. M., Piatachenko V. Yu. et al. Information-Extreme Machine Learning of On-Board Vehicle Recognition System, *Cybernetics and Systems Analysis*, 2020, Vol 56(4), pp. 534–543. DOI:10.1007/s10559-020-00269-y
24. Dovbysh A. S., Budnyk M. M., Moskalenko V. V. Information-Extreme Algorithm for Optimizing Parameters of Hyperellipsoidal Containers of Recognition Classes, *Journal of automation and information sciences*. New York, Begell House Inc., 2012, Vol. 44(10), pp. 35–44. DOI: 10.1615/jautomatInfScien.v44.i10.30.
25. Shelehov I. V., Barchenko N. L., Prylepa D. V. et al. Information-extreme machine training system of functional diagnosis system with hierarchical data structure, *Radio Electronics, Computer Science, Control*, 2022, № 2, pp. 189-200. DOI:10.15588/1607-3274-2022-18.

Received 25.11.2022.

Accepted 22.12.2022.

УДК 681.518:004.93

## ІНФОРМАЦІЙНО-ЕКСТРЕМАЛЬНЕ МАШИННЕ НАВЧАННЯ СИСТЕМИ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ ДЛЯ АДАПТАЦІЇ НАВЧАЛЬНОГО КОНТЕНТУ ДО ВИМОГ РИНКУ ПРАЦІ

**Шелехов І. В.** – доцент, канд. техн. наук, т кафедри комп'ютерних наук, Сумський державний університет, Суми, Україна.

**Прилєпа Д. В.** – асистент кафедри комп'ютерних наук, Сумський державний університет, Суми, Україна.

**Хібовська Ю. О.** – аспірант кафедри комп'ютерних наук, Сумський державний університет, Суми, Україна;

**Отрошенко М. С.** – аспірант кафедри комп'ютерних наук, Сумський державний університет, Суми, Україна.

### АНОТАЦІЯ

**Актуальність.** Розв'язана актуальна задача підвищення функціональної ефективності машинного навчання системи підтримки прийняття рішень (СППР) для оцінки відповідності сучасним вимогам контенту навчальних дисциплін випускової кафедри за результатами опитування роботодавців.

**Мета.** Підвищення функціональної ефективності машинного навчання СППР для оцінки відповідності сучасним вимогам контенту навчальних дисциплін освітньо-професійної програми спеціальності першого (бакалаврського) рівня на основі машинного навчання та розпізнавання образів.

**Метод.** Запропоновано метод інформаційно-екстремального машинного навчання СППР для адаптації навчального контенту випускової кафедри до вимог ринку праці. Ідея методу полягає у максимізації інформаційної спроможності СППР в процесі машинного навчання, що дозволяє в режимі моніторингу гарантувати високу повну ймовірність прийняття правильних класифікаційних рішень. Метод розроблено в рамках функціонального підходу до моделювання когнітивних процесів природнього інтелекту, що дозволяє надати СППР гнучкість при перенавчанні системи через збільшення потужності алфавіту класів розпізнавання. В основу методу покладено принцип максимізації кількості інформації в процесі машинного навчання. Як критерій оптимізації параметрів машинного навчання розглядається модифікована інформаційна міра Кульбака, яка є функціоналом від точнісних характеристик класифікаційних рішень. Згідно із запропонованою функціональною категорійною моделлю розроблено алгоритм інформаційно-екстремального машинного навчання за ієрархічною структурою даних у вигляді бінарного декурсивного дерева. Застосування такої структури даних дозволяє автоматично розбивати велику кількість класів розпізнавання на пари найближчих сусідів, для яких оптимізація параметрів машинного навчання здійснюється за лінійним алгоритмом необхідної глибини. Як параметри оптимізації розглядалися геометричні параметри гіперсферичних контейнерів класів розпізнавання, які в процесі машинного навчання відновлювалися в радіальному базисі бінарного простору ознак Геммінга. При цьому вхідна навчальна матриця трансформувалася в робочу бінарну навчальну матрицю, яка змінювалася в процесі машинного навчання шляхом допустимих перетворень з метою адаптації вхідного інформаційного опису СППР до максимальної достовірності класифікаційних рішень.

**Результати.** Розроблено інформаційне, алгоритмічне і програмне забезпечення СППР для оцінки якості навчального контенту за результатами машинного аналізу відповідей респондентів. В рамках геометричного підходу за результатами інформаційно-екстремального машинного навчання за ієрархічною структурою даних у вигляді бінарного декурсивного дерева побудовано високо достовірні вирішальні правила, практично інваріантні до багато вимірності простору ознак розпізнавання. Досліджено вплив параметрів машинного навчання на функціональну ефективність машинного навчання СППР на прикладі оцінки навчального контенту освітньо-професійної програми бакалаврського рівня спеціальності «Комп'ютерні науки».

**Висновки.** Результати комп'ютерного моделювання підтверджують високу функціональну ефективність запропонованого методу інформаційно-екстремального ієрархічного машинного навчання і можуть бути рекомендовані до

практичного використання у закладах вищої освіти для оцінки відповідності сучасним вимогам навчального контенту випускових кафедр

**КЛЮЧОВІ СЛОВА:** інформаційно-екстремальне машинне навчання, функціональна категорійна модель, інформаційний критерій, ієрархічна структура даних, декурсивне дерево, навчальний контент.

#### ЛІТЕРАТУРА

1. Bates T. National strategies for e-learning in post-secondary education and training / Bates Tony – UNESCO, 2001. – 132 p.
2. Sharable Content Object Reference Model (SCORM) Version 1.2 / Advanced Distributed Learning, 2001. – 55 p.
3. Агарова М. О. Про проблему оптимізації навчального матеріалу / М. О. Агарова // Теорія і практика управління соціальними системами. – 2012. – № 1. – С. 34–39.
4. E-Content: Technologies and Perspectives for the European Market / [eds: A. Nagy, P. Bruck, A. Buchholz, Z. Karszen, A. Zerfass]. – Berlin : Springer-Verlag, 2005. – P. 79–96.
5. Dovgyj S. O. Decision support systems based on statistical and probabilistic procedures / S. O. Dovgyj, P. I. Bidyuk, O. M. Trofymchuk. – Kyiv : Logos, 2014. – 419 p.
6. Distributed optimization and statistical learning via the alternating direction method of multipliers / [S. Boyd, N. Parikh, E. Chu et al.] // Foundations and Trends R in Machine Learning. – 2011. – Vol. 3(1). – P. 1–122.
7. Xu G. Applied Data Mining / G. Xu, Y. Zong, Y. Z. Yang. – CRC Press, 2013. – 284 p.
8. Dua S. Data Mining and Machine Learning in Cybersecurity / S. Dua, X. Du. – Auerbach Publications, 2011. – 256 p.
9. Sklyar V. V. Case-Based and Project-Based Methods for Effective E-learning in ICT Safety and Security / V. V. Sklyar, V. S. Kharchenko // ICTERI Workshops. – 2020. – P. 157–170.
10. Stuart Russell. Intelligence artificielle, 2e éd. (French translation of Artificial Intelligence: A Modern Approach, second edition) / Stuart Russell, Peter Norvig – Paris: Pearson Education France, 2006.
11. Sivanandam S. N. Introduction to fuzzy logic using MATLAB / S. N. Sivanandam, S. Sumathi, S. N. Deepa – Springer, 2007. – 430 p.
12. Watt J. Machine Learning Refined: Foundations, Algorithms, and Applications. – 2nd Edition / J. Watt, R. Borhani, A. Katsaggelos. – Cambridge University Press, 2020. – 594 p.
13. A neural network integrated decision support system for condition-based optimal predictive maintenance policy. Systems / [S.-J. Wu, N. Gebraeel, M. A. Lawley et al.] // Man and Cybernetics, Part A: Systems and Humans. IEEE Transactions. – 2007. – Vol. 37, № 2. – P. 226–236.
14. Radovic M Object recognition in aerial images using convolutional neural networks / M. Radovic, O. Adarkwa, Q. Wang // J. Imaging. – 2017. – Vol. 3, № 2. DOI: 10.3390/jimaging3020021.
15. Alpaydin Ethem. Introduction to Machine Learning / Ethem Alpaydin – London : The MIT Press, 2010. – 110 p.
16. Moskalenko V. V. Information-extreme algorithm of the system for recognition of objects on the terrain with optimization parameter feature extraction / V. V. Moskalenko, A. G. Korobov, // Radio Electronics, Computer Science, Control, 2017. – №2. – P. 38–45.
17. Shi Y. An improvement of neuro-fuzzy learning algorithm for tuning fuzzy rules / Y. Shi, M. Mizumoto // Fuzzy sets and systems. – 2001. – Vol. 118, № 2. – P. 339–350.
18. Subbotin S. A. The neuro-fuzzy network synthesis and simplification on precedents in problems of diagnosis and pattern recognition / S. A. Subbotin // Optical Memory and Neural Networks (Information Optics). – 2013. – Vol. 22, № 2. – P. 97–103. DOI: 10.3103/s1060992x13020082.
19. Efendigil T. A decision support system for demand forecasting with artificial neural networks and neuro-fuzzy models: a comparative analysis / T. Efendigil, S. Önut, C. Kahraman // Expert Systems with Applications. – 2009. – Vol. 36, № 3. – P. 6697–6707.
20. Rahimi A. Random features for large-scale kernel machines / A. Rahimi, B. Recht // Advances in Neural Information Processing Systems. – 2008. – P. 1177–1184.
21. Chapelle O. Training a support vector machine in the primal / O. Chapelle // Neural Computation. – 2007. – Vol. 19, №5. – P. 1155–1178.
22. Ivanciuc O. Applications of Support Vector Machines in Chemistry / O. Ivanciuc // Reviews in Computational Chemistry. – 2007. – Vol. 23. – P. 291–400.
23. Information-Extreme Machine Learning of On-Board Vehicle Recognition System / [A. S. Dovbysh, M. M. Budnyk, V. Yu. Piatachenko et al.] // Cybernetics and Systems Analysis. – 2020. – Vol 56(4). – P. 534–543. DOI:10.1007/s10559-020-00269-y
24. Dovbysh A. S. Information-Extreme Algorithm for Optimizing Parameters of Hyperellipsoidal Containers of Recognition Classes / A. S. Dovbysh, M. M. Budnyk, V. V. Moskalenko // Journal of automation and information sciences. – New York: Begell House Inc. – 2012. – Vol. 44(10) – P. 35–44. DOI: 10.1615/jautomatInfScien.v44.i10.30.
25. Information-extreme machine training system of functional diagnosis system with hierarchical data structure / [I. V. Shelehov, N. L. Barchenko, D. V. Prylepa et al.] // Radio Electronics, Computer Science, Control. – 2022. – № 2. – P. 189–200. DOI:10.15588/1607-3274-2022-18.



## КЛАСТЕРИЗАЦІЯ МАСИВІВ ДАНИХ НА ОСНОВІ МОДИФІКОВАНОГО АЛГОРИТМУ СІРОГО ВОВКА

**Шафроненко А. Ю.** – канд. техн. наук, доцент, доцент кафедри інформатики, Харківський національний університет радіоелектроніки, Харків, Україна.

**Бодяньський С. В.** – д-р техн. наук, професор, професор кафедри штучного інтелекту, Харківський національний університет радіоелектроніки, Харків, Україна.

**Головін О. О.** – д-р техн. наук, с.н.с., заступник начальника Центрального науково-дослідного інституту озброєння та військової техніки Збройних Сил України, Київ, Україна.

### АНОТАЦІЯ

**Актуальність.** Задача кластеризації масивів багатовимірних даних, основною метою якої є знаходження однорідних у сенсі прийнятої метрики класів спостережень, є важливою частиною інтелектуального аналізу даних Data Mining. З обчислювальної точки зору задача кластеризації перетворюється у проблему пошуку локальних екстремумів багатоекстремальної функції, які багатократно запускаються з різних точок вихідного масиву даних. Пришвидшити процес пошуку цих екстремумів можна, скориставшись ідеями еволюційної оптимізації, що включає в себе алгоритми, інспіровані природою, ройові алгоритми, популяційні алгоритми, тощо.

**Мета.** Мета роботи полягає у запровадженні процедури кластеризації масивів даних на основі покращеного алгоритму сірого вовка.

**Метод.** Введено метод кластеризації масивів даних на основі модифікованого алгоритму сірого вовка. Перевагою запропонованого підходу є скорочення часу вирішення оптимізаційних задач в умовах коли кластери перетинаються. Особливістю запропонованого методу є обчислювальна простота і висока швидкість, пов'язана з тим, що весь масив обробляється тільки один раз, тобто виключається необхідність в багатоепоховому самонавчанні, що реалізується в традиційних алгоритмах нечіткої кластеризації.

**Результати.** Результати експериментів підтверджують ефективність запропонованого підходу в задачах кластеризації за умов перетинних кластерів та дозволяють рекомендувати запропонований метод для використання на практиці для вирішення проблем автоматичної кластеризації великих даних.

**Висновки.** Введено метод кластеризації масивів даних на основі покращеного алгоритму сірого вовка. Перевагою запропонованого підходу є скорочення часу вирішення оптимізаційних задач. Результати експериментів підтверджують ефективність запропонованого підходу в задачах кластеризації за умов перетинних кластерів.

**КЛЮЧОВІ СЛОВА:** нечітка кластеризація, багатоекстремальна оптимізація, еволюційний метод.

### АБРЕВІАТУРИ

FCM – метод нечітких  $c$ -середніх;  
HSCI – гібридні системи обчислювального інтелекту;  
PCO – алгоритм «рій частинок»;  
CSO – алгоритм зграї котів;  
NIC – природньо – інспіровані обчислення;  
GWO – алгоритм сірого вовка.

### НОМЕНКЛАТУРА

$X$  – матриця набору даних;  
 $E$  – цільова функція;  
 $N$  – кількість спостережень;  
 $R$  – вектор атрибутів;  
 $n$  – кількість атрибутів;  
 $k$  – номер вектору-спостереження;  
 $x(k)$  – вектор-спостереження;  
 $x_i(k)$  – значення вектора-спостереження за  $i$ -м атрибутом;  
 $x_{i,i2}(k)$  – значення вектора-спостереження за  $i$ -м та  $i2$ -м атрибутами при нормуванні даних в гіперкуб;  
 $i$  – номер атрибуту вектора-спостереження;  
 $j$  – номер кластеру;

$m$  – кількість неперетинних класів;  
 $w$  – вага вовка;  
 $c$  – центроїд кластера;  
 $c_j$  – центроїд  $j$ -го кластеру;  
 $t$  – ітерація пошуку;  
 $T$  – максимальна кількість ітерацій, що задана;  
 $\phi$  – фаззіфікатор;  
 $\alpha$  – контрольний параметр;  
 $r$  – випадкове число;  
 $U$  – рівень належності спостереження до кластеру;  
 $U_j$  – рівень належності спостереження до  $j$  кластеру;  
 $U_j(k)$  – рівень належності  $k$ -го вектора – спостереження до  $j$ -го кластера;  
 $U_j^\phi(k)$  – рівень належності  $k$ -го вектора-спостереження до  $j$ -го кластера при заданому рівні розмитості (фаззіфікатора)  $\phi$  кластерів, що перетинаються;  
 $A, B, C$  – коефіцієнти поведінки оточення;  
 $GW$  – вектор позиції сірого вовка;

$GW(t)$  – вектор позиції сірого вовка в поточній ітерації  $t$ ;

$\alpha$  -,  $\beta$  - та  $\delta$  –вовки-домінанти.

## ВСТУП

Задача кластеризації масивів багатовимірних даних, основною метою якої є знаходження однорідних у сенсі прийнятої метрики класів спостережень, є важливою частиною інтелектуального аналізу даних Data Mining [1–3]. В рамках традиційного кластерного аналізу апріорі передбачається, що кожен вектор спостереження може належати тільки одному класу-кластеру, хоча в реальних даних досить часто виникає ситуація, коли це спостереження з різними рівнями належності (можливості, ймовірності) відноситься відразу до кількох кластерів, що взаємно перетинаються. Подібна ситуація є предметом розгляду нечіткого (фаззі -) кластерного аналізу [4–5], в рамках якого необхідно оцінити не тільки факт належності кожного спостереження до конкретних класів, але і дати кількісну оцінку рівня цієї належності. З обчислювальної точки зору можна відзначити, що найбільш адекватним математичним апаратом для вирішення задач кластеризації є методи штучного інтелекту [6–8] і, перш за все, нейронні мережі, нечіткі системи, еволюційна оптимізація та, так звані, гібридні системи обчислювального інтелекту.

**Об’єкт дослідження** кластеризація даних на основі покращеного алгоритму сірого вовка.

**Предмет дослідження** процедура онлайн кластеризації даних в умовах кластерів що перетинаються на основі модернізованої еволюційної оптимізації.

**Мета роботи** полягає у запровадженні процедури кластеризації масивів даних на основі покращеного алгоритму сірого вовка.

## 1 ПОСТАНОВКА ЗАВДАННЯ

Вихідною інформацією для вирішення задачі кластеризації традиційно є матриця спостережень

$$X = \{x(1), x(2), \dots, x(k), \dots, x(N)\},$$

$x(k) = \{x_i(k)\} \in R^n$ , при цьому дані попередньо відцентровано на гіперкуб так, що  $x(k) = \{x_{i_1, i_2}(k)\} \in R^{n_1 \times n_2}$ . Така ситуація може виникати у випадку обробки масивів зображень.

## 2 ОГЛЯД ЛІТЕРАТУРИ

На сьогодні крім методів нечіткої кластеризації таких як FCM, розроблено безліч методів і алгоритмів нечіткої класифікації зі своїми достоїнствами і недоліками, всі вони дозволяють відшукати тільки локальний екстремум прийнятої цільової функції [5, 9], що веде до того, що використання процедур оптимізації (нелінійного програмування) на основі похідних прийнятого критерію в загальному випадку не дозволяє

отримати найкраще шукане рішення. Подолати цю проблему можна, багаторазово вирішуючи задачу за різних початкових умов і вибираючи найкращий варіант із безлічі отриманих. Зрозуміло, що подібний підхід суттєво збільшує час вирішення задачі.

Подолати зазначені труднощі можна, скориставшись апаратом гібридних систем обчислювального інтелекту (HSCI) [7, 9–11], що поєднують в собі навчання штучних нейронних мереж, інтерпретованість результатів і можливість роботи в умовах класів, що перетинаються, систем непарного виведення і високу швидкість відшукування глобального екстремуму, що забезпечується еволюційними алгоритмами оптимізації, заснованими на «роях частинок» (PCO).

Традиційно алгоритми оптимізації поділяються на дві частини: детерміновані алгоритми та стохастичні алгоритми [13]. Доведено, що детерміновані алгоритми легко потрапляють в локальні оптимуми, в той час як стохастичні алгоритми здатні уникати локальних розв’язків випадковим чином. Таким чином стохастичні алгоритми набули широкого розвитку, зокрема презентацій, удосконалень і застосувань природно-інспірованих обчислень (NIC).

Однією з найважливіших частин алгоритмів NIC є так звані біонічні алгоритми, і більшість яких є метаевристичними [13–15]. Вони можуть вирішувати проблеми з паралельними обчисленнями та глобальним пошуком. Метаевристичні алгоритми поділяють рої на глобальний і локальний пошук за допомогою деяких методів. NIC алгоритми не можуть гарантувати глобальні оптимальні рішення; таким чином, більшість метаевристичних алгоритмів вводять випадковість, щоб уникнути локальних оптимумів. Індивідуумами в зграях керують, щоб розділяти, вирівнювати та об’єднувати за допомогою випадковості; їх поточні швидкості складаються з попередніх швидкостей, випадкових множників частоти [16] або евклідових відстаней положень конкретних індивідів [17–21]. Деякі покращення зроблено за допомогою модифікації ваг інерції, хаосу та бінарних векторів, тощо. Більшість із цих удосконалень призводить до трохи кращої продуктивності конкретних алгоритмів, але загальні структури залишаються незмінними.

Більшість метаевристичних алгоритмів та їх удосконалення наразі базуються безпосередньо на поведінці організмів, таких як пошук, полювання [18], запилення [19] та спалах [20].

Метаевристичні алгоритми працюють за схожими цільовими функціями, та досягають кращої продуктивності та зменшення ймовірності потрапити в пастку локальних оптимумів, уникнути випадкових блукань або польотів за допомогою введення додаткових умов для індивідів. Здебільшого це означає, що зграї поведуться більш неконтрольованими способами. Крім того, як організми, що живуть у зграях в природі, більшість із них мають соціальну ієрархію. Наприклад, у мурашиній колонії королева є командиром, незважаючи на її репродуктивну роль; динергати – це солдати, які займаються садівництвом колонії, тоді як

ергати займаються будівництвом, збиранням і розведенням.

### 3 МАТЕРІАЛИ І МЕТОДИ

В основі поширеного алгоритму ймовірної нечіткої кластеризації лежить процедура мінімізації цільової функції

$$E(U_j(k), c_j) = \sum_{k=1}^N \sum_{j=1}^m U_j^\varphi(k) \|x(k) - c_j\|^2 \quad (1)$$

при обмеженнях

$$\sum_{j=1}^m U_j(k) = 1, \quad 0 \leq \sum_{j=1}^m U_j(k) \leq N, \quad (2)$$

(тут  $\varphi$  – невід’ємний параметр фазифікації (фазифікатор), що задає розмитість границь між кластерами), в основі якого лежать стандартні методи нелінійного (при  $\varphi = 2$  – квадратичного) програмування.

Вирішуючи задачу нелінійного програмування, отримуємо імовірнісний алгоритм нечіткої кластеризації

$$\begin{cases} U_j(k) = \frac{\left(\|x(k) - c_j\|^2\right)^{-1}}{\sum_{l=1}^m \left(\|x(k) - c_l\|^2\right)^{-1}}, \\ c_j = \frac{\sum_{k=1}^N U_j^2(k) x(k)}{\sum_{k=1}^N U_j^2(k)}. \end{cases} \quad (3)$$

В [5] була показана збіжність процесу (3) до локального мінімуму, при цьому досягнення глобального екстремуму в загальному випадку не гарантується.

В роботах [22–23] задача умовної оптимізації (1), (2) була переформульована в задачу безумовної оптимізації цільової функції виду

$$E(c_j) = \sum_{k=1}^N \left( \sum_{j=1}^m \|x(k) - c_j\|^{2(1-\varphi)} \right)^{1-\varphi}, \quad (4)$$

при  $\varphi = 2$

$$E(c_j) = \sum_{k=1}^N \left( \sum_{j=1}^m \|x(k) - c_j\|^{-2} \right)^{-1}. \quad (5)$$

Таким чином, задача нечіткої кластеризації може бути зведена до пошуку глобального екстремуму цільових функцій (4), (5).

Для вирішення задачі можуть бути використані еволюційні біоінспіровані «роєві» процедури оптимізації [9–11], серед яких в якості одного з найбільш швидкодіючих можна відзначити, так званий, алгоритм сірого вовка (GWO) [24].

За даними Мірджалілі [24], сірі вовки живуть разом і полюють групами. Процес пошуку та полювання можна описати так: (6) якщо видобуток знайдено, вони спочатку вистежують, переслідують і наближаються до неї; (7) якщо здобич біжить, тоді сірі вовки переслідують, оточують і спостерігають за здобиччю, поки вона не перестане рухатися; (8) нарешті починається атака.

Стандартний алгоритм GWO. Алгоритм імітує поведінку пошуку і полювання на здобич сірих вовків в зграї. В математичній моделі найкращий результат вовка в зграї називається альфа ( $\alpha$ ), а другий найкращий – бета ( $\beta$ ), і, отже, третій найкращий називається дельта ( $\delta$ ). Інші рішення кандидатів зграї омегами ( $\omega$ ). Всі омеги будуть керуватися цими трьома сірими вовками під час пошуку (оптимізації) та полювання.

Коли жертва знайдена, починається ітерація ( $t=1$ ). Згодом  $\alpha$ -,  $\beta$ - та  $\delta$ -вовки керуватимуть  $\omega$ , щоб переслідувати здобич і, зрештою, оточити її. Три коефіцієнти  $A$ ,  $B$  і  $C$  пропонуються для опису поведінки оточення:

$$\begin{aligned} C_\alpha &= |B_1 * GW_\alpha - X(t)|, \\ C_\beta &= |B_2 * GW_\beta - X(t)|, \\ C_\delta &= |B_3 * GW_\delta - X(t)|, \end{aligned} \quad (6)$$

де  $t$  вказує на поточну ітерацію,  $GW$  вектор позиції сірого вовка,  $GW_1, GW_2$  і  $GW_3$  – є векторами положення  $\alpha$ -,  $\beta$ - та  $\delta$ -вовків, що обчислюється наступним чином:

$$\begin{aligned} GW_1 &= GW_\alpha - A_1 * C_\alpha, \\ GW_2 &= GW_\beta - A_2 * C_\beta, \\ GW_3 &= GW_\delta - A_3 * C_\delta, \end{aligned} \quad (7)$$

$$GW(t) = \frac{GW_1 + GW_2 + GW_3}{3}. \quad (8)$$

Параметри  $A$  та  $B$  є комбінаціями керуючого параметра  $a$  та випадкових чисел  $r_1$  та  $r_2$  [24]:

$$\begin{aligned} A &= 2ar_1 - a, \\ B &= 2r_2. \end{aligned} \quad (9)$$

Контрольний параметр  $a$  замінюється значенням параметра  $A$  і, нарешті, змушує омега-вовків наближатися або тікати від домінуючих вовків, таких як альфа, бета та дельта. Якщо  $|A| > 1$ , сірі вовки втікають від домінантів, а це означає, що омега-вовки вте-

чуть від здобичі та досліджуватимуть більше простору, що в оптимізації називається глобальним пошуком. Та якщо  $|A| < 1$  вони наближаються до домінант, а значить  $\delta$ -вовки будуть слідувати за домінантами, які наближаються до здобичі, і це називається локальним пошуком в оптимізації.

Контрольний параметр  $\alpha$  визначається як лінійне зниження від максимального значення 2 до 0 під час ітерацій:

$$\alpha = 2 \left( 1 - \frac{t}{T} \right),$$

де  $t$  – номер ітерації,  $T$  – максимальна кількість ітерацій, що задана.

Схематично представити роботу алгоритму можна наступним чином (рис. 1).

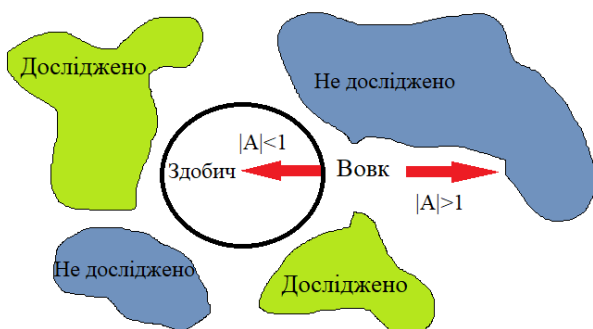


Рисунок 1 – Схема роботи алгоритму GWO

Блок-схема алгоритму сірих вовків наведена на рис. 2.

Багато алгоритмів ройового інтелекту імітують поведінку полювання та пошуку деяких тварин. Однак GWO моделює внутрішню ієрархію керівництва вовків, таким чином, в процесі пошуку позиція найкращого рішення може бути комплексно оцінена рьома рішеннями. Але для інших алгоритмів ройового інтелекту, найкраще рішення шукається лише на основі одного рішення – локального оптимума.

Отже, GWO може значно зменшити ймовірність передчасного потрапляння в локальний оптимум. Щоб досягти належного компромісу між розвідкою та полюванням, пропонується покращений GWO.

Розглядаючи рівняння (8) видно, що в процесі пошуку, однакову роль відіграють домінанти. Кожен із сірих вовків зграї наближається або тікає в пошуку здобичі. Однак, слід зауважити, що найближче до здобичі домінанти із середньою вагою альфа, ніж бета і дельта. Таким чином, на початку процедури пошуку в рівнянні (8) слід враховувати лише положення альфа, або його вага має бути набагато більшою, ніж ваги інших домінант. Таким чином, рівняння (8) можна переписати у вигляді:

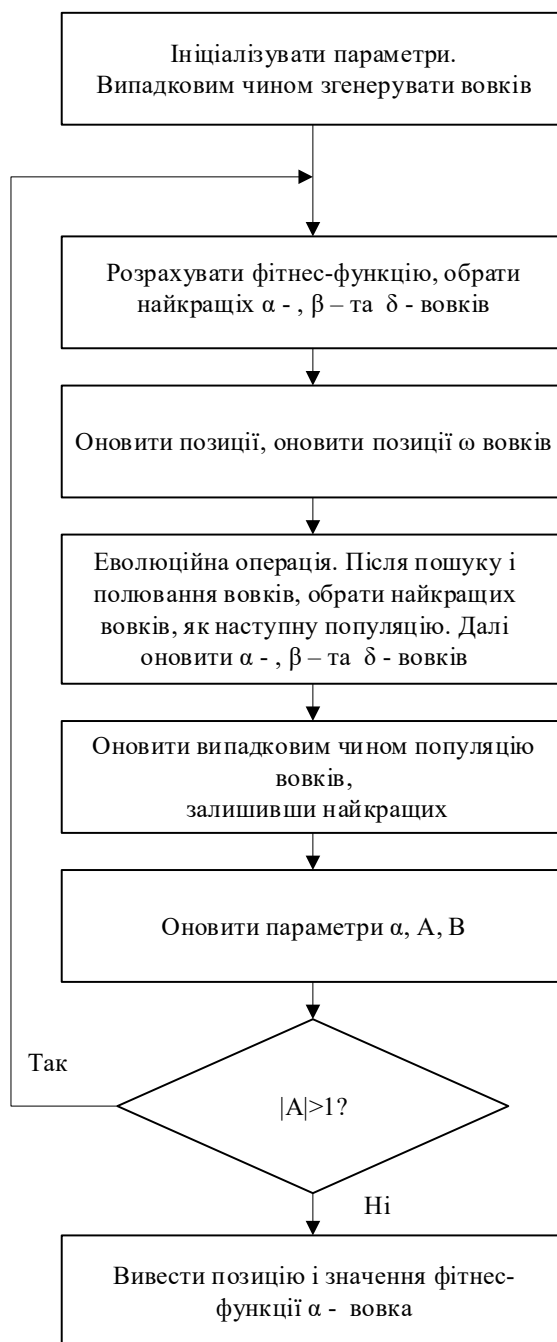


Рисунок 2 – Блок-схема алгоритму GWO

$$GW(t+1) = \frac{w_1 GW_1 + w_2 GW_2 + w_3 GW_3}{3}, \quad (10)$$

де  $w_1 + w_2 + w_3 = 1$ , при  $w_1$  – вага  $\alpha$ -вовка,  $w_2$  – вага  $\beta$ -вовка,  $w_3$  – вага  $\delta$ -вовка, при цьому  $w_1 \geq w_2 \geq w_3$ . На першій (або  $t = 0$ ) ітерації пропонується задати ваги результатами алгоритму кластеризації за рівнянням (3), де:

$$\begin{aligned} c_1 &= w_\alpha; \\ c_2 &= w_\beta; \text{ при } t = 0 \\ c_3 &= w_\delta; \end{aligned} \quad (11)$$

Тоді ми можемо визначити, що ваги змінних задовольняють гіпотезі про соціальну ієрархію функцій сірих вовків та їх пошукову поведінку.

#### 4 ЕКСПЕРИМЕНТИ

Дослідження методу кластеризації масивів даних на основі покращеного алгоритму сірого вовка (FGWO) проводились на двох багатоекстремальних функціях, наведених в табл. 1.

Таблиця 1 – Тестові функції

Назва функції	Формула	Інтервал
Растрігін	$f(x) = 20 + x^2 + y^2 - 10 \cos(2\pi x) + \cos(2\pi y)$	$[-5.12; 5.12]$
Гриванг	$f(x) = \frac{1}{4000}x + \frac{1}{4000}y - \cos\left(\frac{x}{\sqrt{1}}\right)\cos\left(\frac{y}{\sqrt{2}}\right) + 1$	$[-30; 30]$

Якість роботи запропонованого методу (FGWO) порівнювалось із декількома класичними алгоритмами кластеризації, еволюційними процедурами, а також модифікованими методами кластеризації на основі оптимізаційних процедур, а саме алгоритм оптимізації рою частинок (PSO), алгоритм зграї котів (CSO), класичний алгоритм сірого вовка (GWO) та модифікованого алгоритму кластеризації на основі зграї котів (FCSO) [25–26]. Для кожного метода, задано 30 агентів, що шукають оптимум в багатоекстремальній функції.

#### 5 РЕЗУЛЬТАТИ

Перш за все перевіримо роботу запропонованого метода з його модифікацією, тобто використання вагів для кожного вовка. Результат зміни ваг продемонстровано на Рисунку 3. Аналізуючи отриманий графік залежності зміни ваг кожного вовка від кількості ітерацій, можна зробити висновок, що запропонований підхід є сприятливим для подальшого аналізу методу кластеризації масивів даних на основі покращеного алгоритму сірого вовка.

На рис. 4 та рис. 5 показане графічне порівняння методів та їх збіжності за функціями Растрігін та Гриварга відповідно.

#### 6 ОБГОВОРЕННЯ

Аналізуючи результати отриманих експериментальних досліджень та порівняльного аналізу роботи методу кластеризації масивів даних на основі покращеного алгоритму сірого вовка із методами кластеризації, що базуються як на класичному підході до кластеризації даних, так і більш екзотичних, запропонований метод демонструє достатньо високі результати.

Основними перевагами запропонованого методу полягає в простоті математичних розрахунків, швидкості роботи з даними, незалежно від виду, розміру та якості вибірки, що аналізується. Слід відзначити точність роботи метода кластеризації даних на основі по-

кращеного алгоритму сірого та отриманих результатів кластеризації, що досягається за допомогою оптимізаційної процедури еволюційного алгоритму.

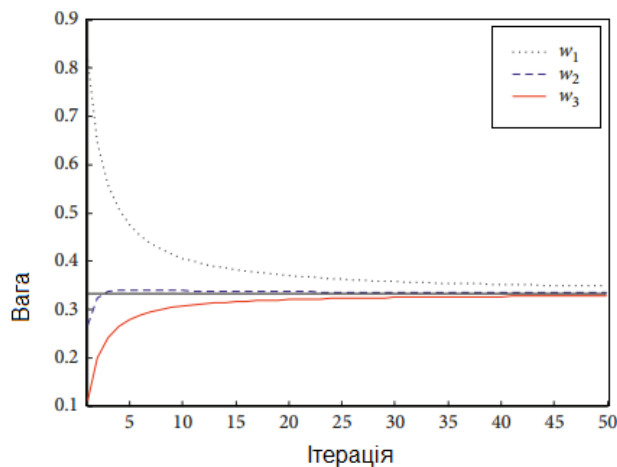


Рисунок 3 – Залежність зміни ваги вовків від кількості ітерацій

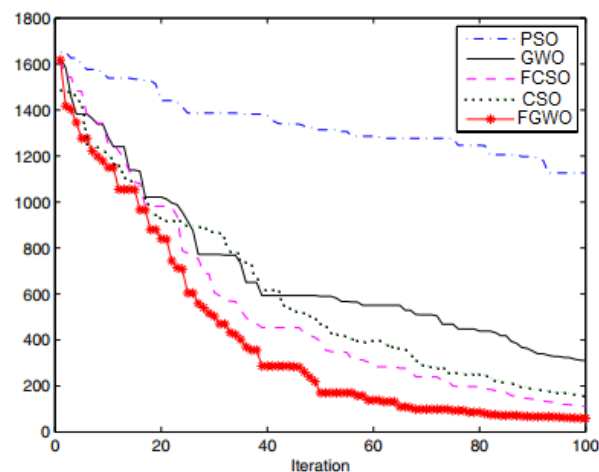


Рисунок 4 – Криві збіжності функції Растрігін

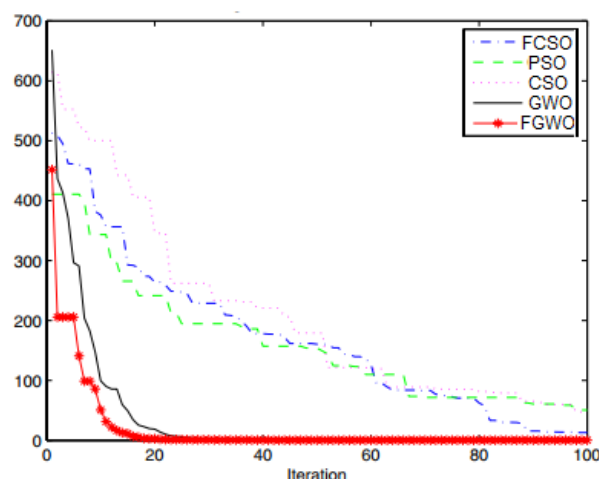


Рисунок 5 – Криві збіжності функції Гриванга

## ВИСНОВКИ

Введено метод кластеризації масивів даних на основі покращеного алгоритму сірого вовка. Перевагою запропонованого підходу є скорочення часу вирішення оптимізаційних задач. Результати експериментів підтверджують ефективність запропонованого підходу в задачах кластеризації за умов перетинних кластерів.

**Наукова новизна:** вперше запропонований метод кластеризації масивів даних на основі покращеного алгоритму сірого вовка.

**Практичне значення:** результати експерименту дозволяють рекомендувати запропонований метод для використання на практиці для вирішення проблем автоматичної кластеризації багатоекстремальних даних різної природи.

**Перспективи подальших досліджень** методи нечіткої кластеризації даних для широкого класу практичних проблем.

## ПОДЯКА

Робота виконана в рамках науково-дослідного проєкту державного бюджету Харківського національного університету радіоелектроніки «Розробка методів та алгоритмів комбінованого навчання глибоких нейро-нео-фаззі систем за умов короткої навчальної вибірки».

## ЛІТЕРАТУРА

1. Gan G. Data Clustering: Theory, Algorithms and Applications/ G. Gan, Ch. Ma, J. Wu. – Philadelphia, Pennsylvania: SIAM: 2007. – 455 p. doi: <https://doi.org/10.1137/1.9780898718348>
2. Abonyi J. Cluster Analysis for Data Mining and System Identification / J. Abonyi, D. Feil. – Basel : Birlhause, 2007 – 303p.
3. Xu R. Clustering/ R. Xu, D. C. Wunsch. – Hoboken N.J. : John Wiley & Sons, Inc., 2009. – 398p.
4. Fuzzy Clustering Analysis: Methods for Classification, Data Analysis and Image Recognition / [Höppner F., Klawonn F., Kruse R., Runkler T.] – Chichester: John Wiley & Sons, 1999. – 300 p.
5. Fuzzy models and algorithms for pattern recognition and image processing / Bezdek J. C. et al.– Springer Science & Business Media, 1999. – Т. 4.
6. Engelbrecht A. Computational intelligence: an introduction / A. Engelbrecht. – Sidney: John Wiley & Sons, 2007. – 597 p.
7. Rutkowski L. Computational Intelligence Methods and Techniques / L. Rutkowski. - Berlin Heidelberg : Springer-Verlag, 2008. – 514 p.
8. Kroll A. Computational Intelligence. Eine Einführung in Probleme, Methoden and Technische Anwendungen / A. Kroll. – München : Oldenbourg Verlag, 2013. – 428 p.
9. Bezdek J. C. Fuzzy Models and Algorithms for Pattern Recognition and Image Processing / [J. C. Bezdek, J. Keller, R. Krishnapuram, N. R. Pal]. – N.Y. : Springer Science + Business Media, Inc., 2015. – 776 p.
10. Mumford C. L. Computational Intelligence/ C. L. Mumford, L.C. Jain. – Berlin: Springer-Verlag, 2009. – 729 p.
11. Kroll A. Computational Intelligence. Eine Einführung in Probleme, Methoden and Technische Anwendungen / A. Kroll. – München : Oldenbourg Verlag, 2013 – 428 p.
12. Mirjalili S. The ant lion optimizer / S. Mirjalili // Advances in Engineering Software. – 2015. – Vol. 83. – P. 80–98. doi: <https://doi.org/10.1016/j.advengsoft.2015.01.010>.
13. Bio-inspired Computation in Telecommunications / [X. S. Yang, S. F. Chen, and T. O. Ting]. –Morgan Kaufmann, Boston, MA, USA, 2015.
14. Syberfeldt A. Real-world simulation-based manufacturing optimizations using cuckoo search / A. Syberfeldt, S. Lidberg // Proceedings of the 2012 Winter Simulation Conference (WSC): Berlin, Germany, December 2012: proceedings. – P. 1–12. doi: 10.1109/WSC.2012.6465158
15. Coelho L. D. S. Improved firefly algorithm approach applied to chiller loading for energy conservation / L. D. S. Coelho and V. C. Mariani // Energy and Buildings. – 2013. – Vol. 59. – P. 273–278. doi: <https://doi.org/10.1016/j.enbuild.2012.11.030>
16. Juan Z. The Bat Algorithm and Its Parameters, Electronics, Communications and Networks IV / Z. Juan, G. Zheng-Ming. – CRC Press, Boca Raton, FL, USA, 2015.
17. Yu. J. J. Q. A social spider algorithm for global optimization/ J. J. Q. Yu and V. O. K. Li// Applied Soft Computing. – 2015. – Vol. 30. – P. 614–627. doi: <https://doi.org/10.48550/arXiv.1502.02407>
18. Azizi R. Empirical study of artificial fish swarm algorithm / R. Azizi // International Journal of Computing, Communications and Networking. – 2014. – Vol. 3, No. 1–3. – P. 1–7.
19. Yan-Xia L. Improved ant colony algorithm for evaluation of graduates/ L. Yan-Xia, L. Lin, and Zhaoyang//Physical conditions, measuring technology and mechatronics automation (ICMTMA): proceedings of the 2014 Sixth International Conference on Measuring Technology and Mechatronics Automation: Zhangjiajie, China, January 2014. – P. 333–336.
20. A modification of artificial bee colony algorithm applied to loudspeaker design problem/ [Z. Xiu, Z. Xin, S. L. Ho, and W. N. Fu] // IEEE Transactions on Magnetics. – 2014. –Vol. 50, No. 2. – P. 737–740. doi: 10.1109/TMAG.2013.2281818.
21. Marichelvam M. K. A discrete firefly algorithm for the multi-objective hybrid flowshop scheduling problems/ M. K. Marichelvam, T. Prabaharan, and X. S. Yang // IEEE Transactions on Evolutionary Computation. – 2014. – vol. 18, No. 2. – P. 301–305.
22. Hathaway R. J. Optimization of clustering criteria by reformulation/ R. J. Hathaway, J. C Bezdek// IEEE Transactions Fuzzy Systems. – 1995. – No. 3. – P. 241–245.
23. Pal N. R. Sequential competitive learning algorithm / N. R. Pal, J. C. Bezdek, R. J. Hathaway // Neural Networks. – 1996. – Vol. 9, № 5. – P.787–796.
24. Mirjalili S. M. Grey wolf optimizer / S. M. Mirjalili and A. Lewis // Advances in Engineering Software. – 2014. – Vol. 69. – P. 46–61.
25. Бодяньський Є. В. Кластеризація масивів даних на основі комбінованої оптимізації функцій щільності розподілу та еволюційного методу котячих зграй/ Є. В. Бодяньський, І. П. Плісс, А. Ю. Шафроненко // Радіоелектроніка, інформатика, управління. – 2022. – № 4. – С. 61–70. doi: 10.15588/1607-3274-2022-4-5
26. Bodyanskiy Y. V. Credibilistic fuzzy clustering based on evolutionary method of crazy cats / Y. V. Bodyanskiy, A. Y. Shafronenko, I. P. Pliss // System Research and Information Technologies. – 2021 (3). – 2021. – P.110–119.

Стаття надійшла до редакції 09.01.2023.

Після доробки 03.02.2023.

UDC 004.8:004.032.26

## CLUSTERIZATION OF DATA ARRAYS BASED ON THE MODIFIED GRAY WOLF ALGORITHM

**Shafronenko A. Yu.** – PhD, Associated Professor at the Department of Informatics, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

© Шафроненко А. Ю., Бодяньський Є. В., Головін О. О., 2023  
DOI 10.15588/1607-3274-2023-1-7

**Bodyanskiy Ye. V.** – Dr. Sc., Professor at the Department of Artificial Intelligence, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.

**Holovin O. O.** – Dr. Sc., Senior Researcher, Deputy Chief of Central Scientific Research Institute of Armament and Military Equipment of Armed Forces of Ukraine, Kyiv, Ukraine.

#### ABSTRACT

**Context.** The task of clustering arrays of multidimensional data, the main goal of which is to find classes of observations that are homogeneous in the sense of the accepted metric, is an important part of the intelligent data analysis of Data Mining. From a computational point of view, the problem of clustering turns into the problem of finding local extrema of a multiextreme function, which are repeatedly started from different points of the original data array. To speed up the process of finding these extrema using the ideas of evolutionary optimization, which includes algorithms inspired by nature, swarm algorithms, population algorithms, etc.

**Objective.** The purpose of the work is to introduce a procedure for clustering data arrays based on the improved gray wolf algorithm.

**Method.** A method of clustering data arrays based on the modified gray wolf algorithm is introduced. The advantage of the proposed approach is a reduction in the time of solving optimization problems in conditions where clusters are overlap. A feature of the proposed method is computational simplicity and high speed, due to the fact that the entire array is processed only once, that is, eliminates the need for multi-era self-learning, implemented in traditional fuzzy clustering algorithms.

**Results.** The results of the experiments confirm the effectiveness of the proposed approach in clustering problems under the condition of classes that overlap and allow us to recommend the proposed method for use in practice to solve problems of automatic clustering big data.

**Conclusions.** A method of clustering data arrays based on the modified gray wolf algorithm is introduced. The advantage of the proposed approach is the reduction of time for solving optimization problems. The results of the experiments confirm the effectiveness of the proposed approach in clustering problems under the conditions of overlapping clusters.

**KEYWORDS:** fuzzy clustering, multi-extremal optimization, evolutionary method.

#### REFERENCES

1. Gan G., Ma Ch., Wu J. Data Clustering: Theory, Algorithms and Applications. Philadelphia, Pennsylvania, SIAM, 2007, 455 p. DOI: <https://doi.org/10.1137/1.9780898718348>
2. Abonyi J., Feil D. Cluster Analysis for Data Mining and System Identification. Basel, Birkhäuser, 2007, 303 p.
3. Xu R., Wunsch D. C. Clustering. Hoboken N. J., John Wiley & Sons, Inc., 2009, 398 p.
4. 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.
5. Bezdek J. C. et al. Fuzzy models and algorithms for pattern recognition and image processing. Springer Science & Business Media, 1999, T. 4.
6. Engelbrecht A. Computational intelligence: an introduction. Sidney, John Wiley & Sons, 2007, 597 p.
7. Rutkowski L. Computational Intelligence Methods and Techniques. Berlin Heidelberg, Springer-Verlag, 2008, 514 p.
8. Kroll A. Computational Intelligence. Eine Einführung in Probleme, Methoden und Technische Anwendungen. München, Oldenbourg Verlag, 2013, 428 p.
9. Bezdek J. C., Keller J., Krishnapuram R., Pal N. R. Fuzzy Models and Algorithms for Pattern Recognition and Image Processing. N.Y., Springer Science + Business Media, Inc., 2015, 776 p.
10. Mumford C. L., Jain L. C. Computational Intelligence. Berlin, Springer-Verlag, 2009, 729 p.
11. Kroll A. Computational Intelligence. Eine Einführung in Probleme, Methoden und Technische Anwendungen. München, Oldenbourg Verlag, 2013, 428 p.
12. Mirjalili S. The ant lion optimizer, *Advances in Engineering Software*, 2015, Vol. 83, pp. 80–98. doi: <https://doi.org/10.1016/j.advengsoft.2015.01.010>.
13. Yang X. S., Chen S. F., and Ting T. O. Bio-inspired Computation in Telecommunications. Morgan Kaufmann, Boston, MA, USA, 2015.
14. Syberfeldt A., Lidberg S. Real-world simulation-based manufacturing optimizations using cuckoo search, *Proceedings of the 2012 Winter Simulation Conference (WSC)*. Berlin, Germany, December 2012, proceedings, pp. 1–12. doi: [10.1109/WSC.2012.6465158](https://doi.org/10.1109/WSC.2012.6465158)
15. Coelho L. D. S. and Mariani V. C. Improved firefly algorithm approach applied to chiller loading for energy conservation, *Energy and Buildings*, 2013, Vol. 59, pp. 273–278. doi: <https://doi.org/10.1016/j.enbuild.2012.11.030>
16. Juan Z., Zheng-Ming G. The Bat Algorithm and Its Parameters, Electronics, Communications and Networks IV. CRC Press, Boca Raton, FL, USA, 2015.
17. Yu. J. J. Q., Li V. O. K. A social spider algorithm for global optimization, *Applied Soft Computing*, 2015, Vol. 30, pp. 614–627. doi: <https://doi.org/10.48550/arXiv.1502.02407>
18. Azizi R. Empirical study of artificial fish swarm algorithm, *International Journal of Computing, Communications and Networking*, 2014, Vol. 3, No. 1–3, pp. 1–7.
19. Yan-Xia L., Lin L., and Zhaoyang Improved ant colony algorithm for evaluation of graduates, *Physical conditions, measuring technology and mechatronics automation (ICMTMA): proceedings of the 2014 Sixth International Conference on Measuring Technology and Mechatronics Automation*. Zhangjiajie, China, January 2014, pp. 333–336.
20. Xiu Z., Xin Z., Ho S. L., and Fu W. N. A modification of artificial bee colony algorithm applied to loudspeaker design problem, *IEEE Transactions on Magnetics*, 2014, Vol. 50, No. 2, pp. 737–740. doi: [10.1109/TMAG.2013.2281818](https://doi.org/10.1109/TMAG.2013.2281818).
21. Marichelvam M. K., Prabakaran T., and Yang X. S. A discrete firefly algorithm for the multi-objective hybrid flowshop scheduling problems, *IEEE Transactions on Evolutionary Computation*, 2014, Vol. 18, No. 2, pp. 301–305.
22. Hathaway R. J., Bezdek J. C. Optimization of clustering criteria by reformulation, *IEEE Transactions Fuzzy Systems*, 1995, No. 3, P.241–245.
23. Pal N. R., Bezdek J. C., Hathaway R. J. Sequential competitive learning algorithm, *Neural Networks*, 1996, Vol. 9, № 5 pp. 787–796.
24. Mirjalili S. M. and Lewis A. Grey wolf optimizer, *Advances in Engineering Software*, 2014, Vol. 69, pp. 46–61.
25. Bodyanskiy Ye. V., Pliss I. P., Shafronenko A. Yu. Clusterization of data arrays based on combined optimization of distribution density functions and the evolutionary method of cat swarm, *Radio Electronics, Computer Science, Control*, 2022, №4, pp. 61–70. doi: [10.15588/1607-3274-2022-4-5](https://doi.org/10.15588/1607-3274-2022-4-5)
26. Bodyanskiy Y.V. Credibilistic fuzzy clustering based on evolutionary method of crazy cats / Y.V. Bodyanskiy, A.Y. Shafronenko, I.P. Pliss // System Research and Information Technologies. – 2021 (3). – 2021. – P.110–119.

# ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

## PROGRESSIVE INFORMATION TECHNOLOGIES

UDC 004.4'24, 004.896

### PARAMETER-DRIVEN GENERATION OF EVALUATION PROGRAM FOR A NEUROEVOLUTION ALGORITHM ON A BINARY MULTIPLEXER EXAMPLE

**Doroshenko A. Yu.** – Dr. Sc., Professor, Head of the Computing Theory Department, Institute of Software Systems of the National Academy of Sciences of Ukraine, Kyiv, Ukraine.

**Achour I. Z.** – Post-graduate student of the Department of Information Systems and Technologies, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine.

**Yatsenko O. A.** – PhD, Senior Researcher of the Computing Theory Department, Institute of Software Systems of the National Academy of Sciences of Ukraine, Kyiv, Ukraine.

#### ABSTRACT

**Context.** The problem of automated development of evaluation programs for the neuroevolution of augmenting topologies. Neuroevolution algorithms apply mechanisms of mutation, recombination, and selection to find neural networks with behavior that satisfies the conditions of a certain formally defined problem. An example of such a problem is finding a neural network that implements a certain digital logic.

**Objective.** The goal of the work is the automated design and generation of an evaluation program for a sample neuroevolution problem (binary multiplexer).

**Method.** The methods and tools of Glushkov’s algebra of algorithms and hyperscheme algebra are applied for the parameter-driven generation of a neuroevolution evaluation program for a binary multiplexer. Glushkov’s algebra is the basis of the algorithmic language intended for multilevel structural design and documentation of sequential and parallel algorithms and programs in a form close to a natural language. Hyperschemes are high-level parameterized specifications intended for solving a certain class of problems. Setting parameter values and subsequent interpretation of hyperschemes allows obtaining algorithms adapted to specific conditions of their use.

**Results.** The facilities of hyperschemes were implemented in the developed integrated toolkit for the automated design and synthesis of programs. Based on algorithm schemes, the system generates programs in a target programming language. The advantage of the system is the possibility of describing algorithm schemes in a natural-linguistic form. An experiment was conducted consisting in the execution of the generated program for the problem of evaluating a binary multiplexer on a distributed cloud platform. The multiplexer example is included in SharpNEAT, an open-source framework that implements the genetic neuroevolution algorithm NEAT for the .NET platform. The parallel distributed implementation of the SharpNEAT was proposed in the previous work of the authors.

**Conclusions.** The conducted experiments demonstrated the possibility of the developed distributed system to perform evaluations on 64 cloud clients-executors and obtain an increase in 60–100% of the maximum capabilities of a single-processor local implementation.

**KEYWORDS:** algebra of algorithms, automated program design, cloud computing, hyperscheme, neuroevolution, neural network, parallel programming.

#### ABBREVIATIONS

IDS is an Integrated toolkit for software Design and Synthesis;

NEAT is NeuroEvolution of Augmenting Topologies;

SAA is a system of algorithmic algebra;

SharpNEAT is an open-source framework written in C# that implements the genetic neuroevolution algorithm NEAT.

#### NOMENCLATURE

$A$  is a nonterminal operator from set  $R_T$ ;

$A_j$  is an operator from set  $\overline{Op}$ ;

$AHS$  is algebra of hyperschemes;

$e$  is an empty word;

$E_3$  is a three-valued logic;

$E_4$  is a four-valued logic;

$F(A, p)$  is a function that specifies the generation method for operations of  $AHS$  signature;

$GA$  is Glushkov’s algebra (system of algorithmic algebra);

$IS$  is a set of states (an information set) of the operational automaton of the abstract automaton model of a computer;

$\overline{L}$  is a set of states of tape  $\tilde{L}$ ;

$\tilde{L}$  is a tape of operational automaton  $\overline{\Phi}$ ;



$m$  is a number of address inputs of a multiplexer;

$\overline{M}$  is a set of states of operational automaton  $\overline{\Phi}$ ;

$\widetilde{M}$  is a stack of control automaton  $\overline{\Psi}$ ;

$n$  is a number of data inputs of a multiplexer;

$Op$  is a set of operators of  $GA$ ;

$\overline{Op}$  is a set of operators of  $AHS$ ;

$p$  is an element of set  $\overline{P}$ ;

$\overline{P}$  is a set of states (an information set) of the operational automaton of the abstract automaton model of the parameter-driven generator of texts;

$P_0$  is an array length;

$P_1$  is a number of address inputs of a multiplexer (hyperscheme parameter);

$P_2$  is a number of information inputs of a multiplexer (hyperscheme parameter);

$P_3$  is the total number of inputs of a multiplexer (hyperscheme parameter);

$P_q$  is a hyperscheme parameter with number  $q$ ;

$Pr$  is a set of predicates of  $GA$ ;

$\overline{Pr}$  is a set of predicates of  $AHS$ ;

$R_N$  is a set of nonterminal operators of  $AHS$ ;

$R_T$  is a set of terminal operators of  $AHS$ ;

$s_j$  is address input of a multiplexer;

$u_k$  is a predicate from set  $\overline{Pr}$ ;

$x_i$  is a data input of a multiplexer;

$y$  is an output of a multiplexer;

$\eta$  is “not computed” value;

$\mu$  is “undefined” value;

$\overline{\Phi}$  is an operational automaton;

$\overline{\Psi}$  is a control automaton;

$\Omega_{AHS}$  is a signature of operations of  $AHS$ ;

$\Omega_{GA}$  is a signature of operations of  $GA$ ;

$\Omega_1$  is a set of logic operations included in  $\Omega_{GA}$ ;

$\Omega_2$  is a set of operator operations included in  $\Omega_{GA}$ .

## INTRODUCTION

One of the promising directions in the development and research of parallel and distributed computing systems is the construction of software abstractions in the form of algebraic-algorithmic languages and models, which aims to develop architecture- and language-independent programming tools for multiprocessor computing systems and networks. In [1], authors proposed a theory, methodology, and software tools for the automated design of parallel programs based on high-level algebraic formalization and automation of program transformations based on rewriting rules. In particular, an instrumental system of programming automation called the integrated toolkit for software design and synthesis (IDS) was developed. One of the important problems within the algebra-algorithmic approach is increasing the adaptability of programs to the specific conditions of their use. In particular, it can be solved by using the method of

parameter-driven generation of algorithm schemes based on higher-level specifications named hyperschemes.

In this paper, the developed algebra-algorithmic facilities are applied to the field of neuroevolution algorithms. Neuroevolution is a promising approach for solving complex problems of machine learning, the development of artificial neural networks, adaptive control, multi-agent systems, and evolutionary robotics [2]. The main advantage of neuroevolution is that it can be used more widely than supervised learning algorithms, which require a program of correct input-output pairs. Neuroevolution only requires evaluating the performance of the network when performing a task. It uses evolutionary algorithms to train a neural network and belongs to the category of reinforcement learning methods. All evolutionary algorithms develop a set (“population”) of solutions (“individuals”). Individuals are represented by their genotype, which is expressed in the form of a phenotype, with which quality, “adaptability” is associated. There are a large number of neuroevolutionary algorithms, divided into two groups. The first includes algorithms that perform the evolution of weights for a given network topology, the second includes algorithms that, in addition to the evolution of weights, also perform the evolution of the network topology. Evolutionary algorithms manipulate a set of genotypes, which are a representation of a neural network. In a direct coding scheme, the genotype is equivalent to the phenotype, and neurons and connections are directly specified in the genotype. Conversely, in the scheme with indirect coding, the rules and structures for creating a neural network are specified in the genotype.

**The object of study** is the automated development of evolutionary algorithms.

One of the implementations of evolutionary algorithms is SharpNEAT [3], an open-source framework developed in the C# language. It implements the genetic neuroevolution algorithm NEAT (NeuroEvolution of Augmenting Topologies) for the .NET platform. The algorithm uses the evolutionary mechanisms of mutation, recombination, and selection to find neural networks with behavior that satisfies the conditions of a certain formally defined problem. Examples of such problems are controlling the movements of a robot’s limbs, flying a rocket, or finding a neural network that implements a certain digital logic (for example, a multiplexer).

Despite the strengths of the NEAT method, such as the possibility of its application in tasks where it is difficult to choose the cost function and neural network topology, one of the problems is the slow convergence to optimal results, especially in the case of complex environments. The distributed implementation of the NEAT evaluation method was proposed in the previous work of the authors [5]. It allows to radically speed up finding optimal configurations of neural networks in the presence of sufficient computing resources.

**The subject of study** is the automated design and generation of evaluation programs for neuroevolution algorithms.

**The purpose of the work** is to apply algorithm algebra and hyperschemes [1, 6] for the parameter-driven generation of an evaluation program for a sample neuroevolution problem.

Hyperschemes are parameterized specifications intended for solving a certain class of problems. Setting specific values of parameters and subsequent interpretation of hyperschemes allows obtaining algorithms adapted to specific conditions of their use. The generator of algorithms based on hyperschemes is one of the components of the above-mentioned IDS toolkit [1]. Algorithm schemes being designed in the toolkit are presented in Glushkov's system of algorithmic algebra (SAA).

The approach to the parameter-driven generation of programs is illustrated on generating the source code of the evaluation procedure for the binary multiplexer problem example included in SharpNEAT [4]. The results of the execution of multi-threaded and distributed versions of the generated procedure on a multicore processor and a cloud platform are given.

### 1 PROBLEM STATEMENT

The problem consists in designing a high-level parameterized specification in the algebra of hyperschemes [1, 6] that is applied to generate classes of evaluation schemes for a binary multiplexer (Binary MultiplexerEvaluator) example [4] depending on the multiplexer parameters, followed by the automated synthesis of code in C# language for the SharpNEAT framework.

A multiplexer is a device that has several data inputs  $x_i$  ( $i = 0, \dots, n-1$ ), address inputs  $s_j$  ( $j = 0, \dots, m-1$ ), and one output  $y$ . The device transmits a signal from one of the data inputs to the output; at the same time, the selection of the desired input is carried out by applying the appropriate combination of control signals to the address inputs. The number of data inputs  $n$  and the number of address inputs  $m$  are related by the ratio:  $n = 2^m$ . The conditional scheme of the multiplexer with 11 inputs is shown in Fig. 1.

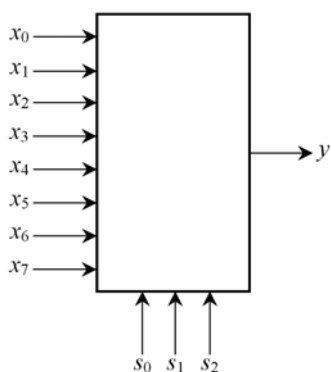


Figure 1 – The conditional scheme of a multiplexer with 11 inputs

The parameters of the hyperscheme are the number  $P_1$  of address inputs of the multiplexer, the number  $P_2$  of its information inputs and the total number of inputs  $P_3 = P_1 + P_2$ . All inputs accept binary values (0 or 1). A binary address is applied to the address inputs, representing the selection of one of the input values for data. The evaluation consists of exhaustively testing the neural network on each of the  $2^{P_3}$  possible input combinations [4]. The output value of the neural network must match the value of one of the data inputs, which is represented by a binary address from the address inputs. An output value less than 0.5 is considered a binary zero, and an output value greater than or equal to 0.5 is a binary one. The value of the assessment (suitability) is calculated additively as a result of the comprehensive check.

Depending on the values of the hyperscheme parameters, a specific scheme of an algorithm in SAA [1] is to be generated, representing a multiplexer evaluation scheme with a specific number of inputs. The examples of parameter values are shown in Table 1. The SAA schemes are the basis for the generation of C# programming code.

Table 1 – The values of the hyperscheme parameters ( $P_1-P_3$ ) for generating multiplexer evaluation schemes

Number of multiplexer inputs	Corresponding values of hyperscheme parameters		
	$P_1$	$P_2$	$P_3$
3	1	2	3
6	2	4	6
11	3	8	11

### 2 REVIEW OF THE LITERATURE

This paper is related to works on the automated generation of programs from specifications [7–10] and neuroevolution of augmenting topologies [2, 11, 12].

In particular, paper [7] presents a tool for the automatic generation of C++ programs from Isabelle (high-order logic theorem prover) specifications. In [8], a combination of code and test generation based on the specification language of the Temporal Logic of Actions (TLA) is proposed. Work [9] presents a tool for generating C++ code from abstract state machine models. Paper [10] proposes an automated technique that generates executable tests from structured natural language specifications.

The peculiar feature of our approach to program generation consists in using algebra of algorithms and hyperschemes [6]. Algorithms and programs are constructed using high-level algebra-algorithmic schemes represented in a natural linguistic form. The developed tools provide automated generation of sequential and parallel code in C++ and Java languages from the schemes. In this paper, we apply these algebra-algorithmic facilities for the automated design of an evaluation procedure for a neuroevolution algorithm.

Neuroevolution of augmenting topologies is a genetic algorithm for finding artificial neural networks through evolution (neuroevolutionary method) [2]. HyperNEAT (Hypercube-based NeuroEvolution of Augmenting To-

pologies) is an extension of NEAT that uses a form of indirect encoding called Compositional Pattern-Producing Networks (CPPNs) [11]. The implementations of NEAT and HyperNEAT are part of a package called SharpNEAT developed in C# by Colin Green [12]. The peculiarity of NEAT and SharpNEAT is that they search both the structure of the neural network (nodes and connections) and the weight parameters of connections between nodes. The parallel distributed version of SharpNEAT was proposed by the authors in [5].

In this work, the distributed version is applied for evaluating the performance of the code generated for binary multiplexer problem example on a cloud platform.

### 3 MATERIALS AND METHODS

In this section, we consider the system of algorithmic algebra and hyperschemes, which are the basis of the algebra-algorithmic approach to algorithm design and synthesis. The software tools for the automated generation of algorithm schemes and programs are also described.

Glushkov's SAA is focused on the analytical form of algorithm representation and formalized transformation of these specifications, in particular, with the aim of optimizing the algorithms according to specified criteria [1]. SAA is the two-sorted algebra  $GA = \langle \{Pr, Op\}; \Omega_{GA} \rangle$ , where sorts are a set  $Pr$  of predicates and a set  $Op$  of operators defined on information set  $IS$ . The operators are mappings (possibly partial) of  $IS$  to itself. The predicates take values of the three-valued logic  $E_3 = \{0, 1, \mu\}$ . The signature  $\Omega_{GA} = \Omega_1 \cup \Omega_2$  consists of system  $\Omega_1$  of logic operations (conjunction, disjunction, negation, and prognosis) that take values in set  $Pr$ , and system  $\Omega_2$  of operator operations (composition, branching, loop, and other) that take values in set  $Op$  and are considered further in more detail.

SAA is the basis of the algorithmic language SAA/1, designed for multilevel structural design and documentation of sequential and parallel algorithms and programs. The advantage of its use is the possibility of describing algorithms in a natural-linguistic form. The operators represented in the SAA/1 language are called SAA schemes. Identifiers of predicates in this language are enclosed in single quotes, and operators – in double ones. Predicates and operators in SAA/1 can be basic or compound. Basic elements are elementary atomic abstractions in algorithm schemes. Compound conditions and operators are built from basic ones using the operations from the SAA signature.

Some operator operations of SAA used in this paper are the following (represented in a natural-linguistic form):

- composition (sequential execution) of operators: “operator1”; “operator2”;
- branching: IF ‘condition’ THEN “operator1” ELSE “operator2” END IF;
- for loop: FOR (counter FROM start TO fin) “operator” END OF LOOP;

© Doroshenko A. Yu., Achour I. Z., Yatsenko O. A., 2023  
DOI 10.15588/1607-3274-2023-1-8

– parallel processing of a list: PARALLEL FOR EACH (elem IN list) (“operator(elem)”).

The algebraic facilities for generation of algorithm schemes are based on SAA and the abstract automaton model of the parameter-driven text generator [1, 6]. The generator works according to a feedback principle (see Fig. 2). The automaton  $\bar{\Psi}$  with stack  $\tilde{M}$  is used as a control automaton, and the automaton  $\bar{\Phi}$  with tape  $\tilde{L}$  is used as an operational one. Tape  $\tilde{L}$  is intended for writing the text of an SAA scheme being generated.

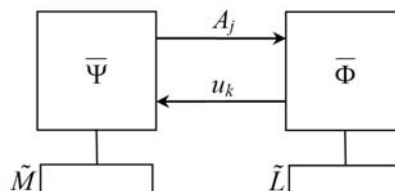


Figure 2 – The abstract automaton model of the parameter-driven text generator

Set  $\bar{M}$  of states of automaton  $\bar{\Phi}$  is associated with parameters that control the generation of schemes. The elements of the information set  $\bar{P} = \bar{M} \times \bar{L}$  are called the states of the operational structure. At each step of the automaton's work, a set of values of logical conditions  $\bar{Pr} = \{u_k\}$  defined on set  $\bar{P}$  is sent from the operational to the control automaton. Depending on these values and contents of stack  $\tilde{M}$ , the control automaton initiates the execution of some operator. The set of operators  $\bar{Op} = \{A_j\}$  is divided into two disjoint sets – terminal operators  $R_T$  and nonterminal operators  $R_N$ . Execution of the terminal operator from set  $R_T$  consists in changing the current state of the operational structure, which, in particular, can be writing some text on tape  $\tilde{L}$ . The execution of operator  $A \in R_N$  at current state  $p \in \bar{P}$  consists in writing some term  $F(A, p)$  to stack  $\tilde{M}$  and its further interpretation by the control automaton. The term  $F(A, p)$  is an analog of the concepts of macro definition, procedure, routine, etc. Stack  $\tilde{M}$  is used at processing nested and recursive terms. The generated text is the content of tape  $\tilde{L}$  in the final state of the operational structure

The considered abstract automaton model is matched with the algebra of hyperschemes intended for the formalization of algorithms for the parameter-driven generation of SAA schemes [6]. It is the two-sorted algebra  $AHS = \langle \{\bar{Pr}, \bar{Op}\}; \Omega_{AHS} \rangle$ , where predicates from set  $\bar{Pr}$  are defined on information set  $\bar{P}$  and take values of the four-valued logic  $E_4 = \{0, 1, \mu, \eta\}$ ; operators from set  $\bar{Op}$  are defined on and take values in set  $\bar{P}$ .

The set of predicates is associated with parameters that control the process of SAA scheme generation. The

operations of the signature  $\Omega_{AHS}$  are similar to the SAA operations. The difference from SAA is that the predicates from set  $\overline{Pr}$  map information set  $\overline{P}$  to set  $E_4$  with additional value  $\eta$ , which is used to indicate that the value of a predicate cannot be computed due to a lack of information about the values of hyperscheme parameters.

The application of operator  $A \in \overline{Op}$  at state  $p \in \overline{P}$  leads to the transition of operational structure  $\overline{\Phi}$  to a new state  $A(p) \in \overline{P}$  and writing some (possibly empty) fragment  $F(A, p)$  of a scheme being generated to tape  $\tilde{L}$ . The function  $F(A, p)$  specifies the generation method for all operations of the algebra of hyperschemes and is defined in detail in [6].

In particular, function  $F(A, p)$  for operation “operator1”; “operator2” generates the composition operation without changes.

For the operation of branching, the generation function is

$$F(A, p) = \begin{cases} \text{"operator1", if 'condition'=1;} \\ \text{"operator 2", if 'condition'=0;} \\ \text{IF 'condition' THEN "operator1" ELSE} \\ \text{"operator2" END, if 'condition'=\eta;} \\ e, \text{if 'condition'=\mu,} \end{cases}$$

where  $e$  is an empty word.

The result of the interpretation of this operation is the text of the first operator at the true value of the condition, and the text of the second operator at the false value. The whole text of the branch operation is generated at a not computed value of the condition. An empty text is a result in the case if there was an error during the interpretation process.

Representations of operators in  $AHS$  are called hyperschemes. Each hyperscheme  $A$  applied at state  $p \in \overline{P}$  generates an SAA scheme  $F(A, p)$ . Hyperscheme  $A$  defines the class of SAA schemes  $\{F(A, p) | p \in \overline{P}\}$ .

The processing of basic conditions and operators of a hyperscheme consists in computing expressions with hyperscheme parameters and substituting them into the text of these basic elements.

The considered approach to the generation of algorithm schemes is implemented in the integrated toolkit for software design and synthesis [1]. Hyperschemes are designed in an automated way by detailing the language constructs of the hyperscheme algebra. The constructs are selected from a list and added to the algorithm design tree. At each step of the design process, the system offers a list of algebra operations depending on the type of tree node selected. A hyperscheme is used for further generation of an SAA scheme of an algorithm (see Fig. 3) and synthesis of a program in a target programming language (C, C++, Java, and other).

To facilitate processing, the parameters are written in the text of the basic and other elements of a hyperscheme in the form  $P_q$  ( $q = 0, 1, 2, \dots$ ). Expressions with hyperscheme parameters are enclosed in square or curly brackets.

**Example.** Consider the use of the hyperscheme facilities for designing a fragment of the hybrid sorting algorithm.

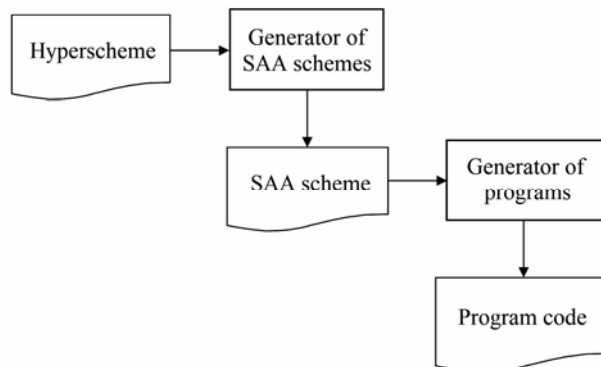


Figure 3 – The sequence of generation of algorithms and programs in the IDS toolkit

In the SAA scheme below, one of the sorting algorithms (insertion, sequential, or parallel merge sort) is selected depending on the length  $P_0$  of the array.

```

“Hybrid sort (array)” ==
= IF ‘[P0 <= 200]’
THEN
  “Insertion sort (array)”
ELSE
  IF ‘[P0 <= 1000]’
  THEN
    “Sequential merge sort (array)”
  ELSE
    “Parallel merge sort (array)”
  END IF
END IF
  
```

Let it be known in advance that the algorithm represented by the scheme will be applied in conditions when  $P_0 \geq 500$ , then the given SAA scheme becomes redundant. Considering it as a hyperscheme, we can assume that at the stage of generation of an SAA scheme, the condition ‘[P0 <= 200]’ takes the value “false”, while ‘[P0 <= 1000]’ takes the value “not computed”. As a result of the generation of text according to the hyperscheme, we will get the shortened SAA scheme:

```

“Hybrid sort (array)” ==
= IF ‘[P0 <= 1000]’ THEN
  “Sequential merge sort (array)”
ELSE
  “Parallel merge sort (array)”
END IF
  
```

#### 4 EXPERIMENTS

In this paper, we apply the facilities of hyperschemes for generating classes of SAA schemes intended for the evaluation of a binary multiplexer (BinaryMultiplexer Evaluator) [4].

The hyperscheme constructed using the IDS toolkit is given below. Its parameters  $P_1, P_2, P_3$  were described in Section 1. In the scheme, curly brackets  $\{P_3\}$  indicate the parameter that needs to be replaced with the corresponding number written in words, that is, if the value of  $P_3 = 11$ , the text “Eleven” will be inserted. Square brackets (for example,  $[P_1]$  or  $[P_3 - 1]$ ) indicate parameters or arithmetic expressions to be replaced by the corresponding number. So, for the loop FOR (i FROM 0 TO  $[Pow(2, P_3) - 1]$ ) at the value of the parameter  $P_3 = 11$ , the text FOR (i FROM 0 TO 2047) will be generated.

SCHEME

```
BINARY {P3} MULTIPLEXEREVALUATOR ==
  “Binary {P3}-multiplexer evaluator scheme”
  END OF COMMENTS
```

```
“Binary {P3} Multiplexer Evaluator” ==
= NAME SPACE
```

```
SharpNeat.Domains.Binary {P3} Multiplexer
(
  CLASS Binary {P3} Multiplexer Evaluator OF
  TYPE public INHERITS
  IPhenomeEvaluator < IBlackBox >
```

```
  “Declare a constant (StopFitness) of type
  (double) = (10E + [P1]);”
  “Declare a variable (_evalCount) of type
  (ulong);”
  “Declare a variable (_stopConditionSatisfied)
  of type (bool);”
```

```
  REGION IPhenomeEvaluator < IBlackBox >
    Members
```

```
  PROPERTY public ulong EvaluationCount
  GET
  (
    “Return value (_evalCount)”
  )
  END OF PROPERTY
```

```
  PROPERTY public bool
    StopConditionSatisfied
  GET
  (
    “Return value (_stopConditionSatisfied)”
  )
  END OF PROPERTY
```

```
  METHOD public FitnessInfo
  Evaluate ( IBlackBox box )
  “Declare a variable (fitness) of type
```

```
(double) = (0.0);”
  “Declare a variable (success) of type
  (bool) = (true);”
  “Declare a variable (output) of type
  (double);”
  “Declare a variable (inputArr) of type
  (ISignalArray) = (box.InputSignalArray);”
  “Declare a variable (outputArr) of type
  (ISignalArray) = (box.OutputSignalArray);”
  “Increase (_evalCount) by (1);”
  FOR (i FROM 0 TO [Pow(2, P3) - 1])
  LOOP
    “Declare a variable (tmp) of type
    (int) = (i);”
    FOR (j FROM 0 TO [P3 - 1])
    LOOP
      (inputArr[j] := tmp & 0x1);
      (tmp := tmp >> 1)
    END OF LOOP;
    “Activate the black box (box);”
    “Read output signal (output)(outputArr);”
    IF (((1 << ([P1] + (i & 0x[P2 - 1]))) & i) != 0)
    THEN
      (fitness := fitness + 1.0 - ((1.0 - output) *
      (1.0 - output)));
      IF (output < 0.5)
      THEN (success := false)
      END IF
    ELSE
      (fitness := fitness + 1.0 - (output *
      output));
      IF (output >= 0.5)
      THEN (success := false)
      END IF
    END IF;
    “Reset black box state ready for next test
    case (box)”
  END OF LOOP;
  IF success
  THEN (fitness := fitness + 10E + [P1])
  END IF;
  IF (fitness >= StopFitness)
  THEN (_stopConditionSatisfied := true)
  END IF;
  “Return value (new
  FitnessInfo(fitness, fitness))”
  END OF METHOD

  METHOD public void Reset()
  “Empty operator”
  END OF METHOD

  END OF REGION

  END OF CLASS
)
END OF SCHEME
```

Based on the hyperscheme, SAA schemes for evaluating multiplexers with three, six, and 11 inputs were generated using the IDS toolkit. Further, C# program code for the SharpNEAT framework was generated according to the schemes.

The scheme of the parallel multi-threaded evaluation procedure for the multiplexer example, implemented in SharpNEAT, looks like this:

```
METHOD private void
  Evaluate_Caching(IList<TGenome> genomeList)
  PARALLEL FOR EACH (genome IN genomeList)
  (
    "Get (phenome) for (genome)";
    IF (phenome = null)
      THEN "Decode the (phenome) and store a
            reference against the (genome)"
      END IF;
    IF (phenome = null)
      THEN
        "Set (genome) fitness to (0.0)";
        "Set (genome) auxiliary fitness info to (null)"
      ELSE
        "Evaluate (phenome) and get fitness
        (fitnessInfo)";
        "Set (genome) fitness to (fitnessInfo._fitness)";
        "Set (genome) auxiliary fitness info to
        (fitnessInfo._auxFitnessArr)"
      END IF
  )
END OF METHOD
```

In [5], the distributed version of this procedure was developed for execution on a cloud computing platform.

In this work, the experiments with multithreaded and distributed implementations of neuroevolution of augmenting topology were carried out. A multiplexer with 11 inputs was selected as an example.

The following configurations were chosen as the execution environments for single-process and distributed implementation:

1) local environment, Intel Core i9-9900K CPU (3.60 GHz – 5.00 GHz), 8 cores, 16 logic processors, 32.0 GB RAM, one process, 16 threads;

2) local environment, Intel Core i9-9900K CPU (3.60 GHz – 5.00 GHz), 8 cores, 16 logic processors, 32.0 GB RAM, distributed implementation, 16 local clients-executors;

3) cloud environment, 3rd Gen AMD EPYC Amazon EC2 C6a.large, 3.60 GHz, 2 cores, 4.0 GB RAM, up to 12.5 Gbit/s of network bandwidth, and up to 6600 Mbit/s of storage bandwidth, distributed implementation, 16 local clients-executors;

4) the same cloud environment, but with 32 cloud client executors;

5) the same cloud environment but with 64 cloud client executors.

## 5 RESULTS

This section gives the results of executing the multiplexer example in the computing environments described above.

Fig. 4 shows the graph of the dependence of the evaluation speed (the number of evaluations per second) on the generation number for local configurations of the environment.

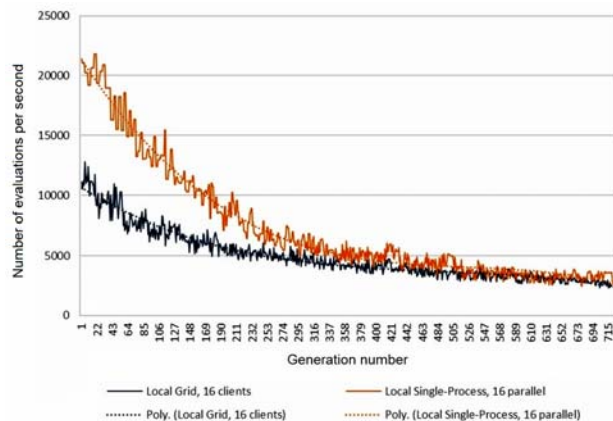


Figure 4 – The graph of the dependence of the evaluation speed on the generation number for local environment configurations

Fig. 5 shows a graph of the evaluation speed on the generation number for the cloud-based environment configurations.

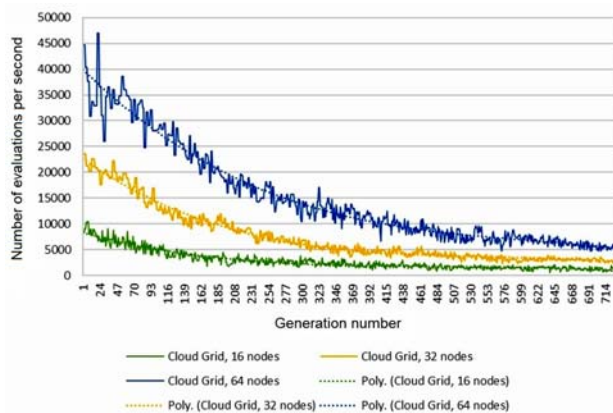


Figure 5 – The graph of the dependence of the evaluation speed on the generation number for cloud environment configurations

## 6 DISCUSSION

As seen from the graph in Fig. 4, the distributed implementation is expected to show worse results compared to the single-process implementation due to the overhead of interaction between processes. As the complexity of the evaluation task increases (the size of the generated neural network increases), the efficiency of the single-process and local distributed implementation is leveled off, since the overhead costs of computing resources become prohibitively lower than the evaluation costs.

As shown in Fig. 5, the distributed cloud implementation is expected to show worse results (for the same number of client-executors) compared to the single-

process and local distributed implementation due to the overhead of interaction between the processors of many computers, clients-executors. However, with the growth of the number of executors, we can neglect the constant value of the overhead and obtain a linear increase in the efficiency of the distributed system.

The results of the experiment demonstrated the ability of the distributed system to conduct evaluations on 64 cloud clients-executors and obtain an increase of 60–100% from the maximum capabilities of a single-processor local implementation.

### CONCLUSIONS

**The scientific novelty** of obtained results is that the facilities of hyperscheme algebra are firstly applied for the automated generation of parametric neuroevolution evaluation algorithms on the example of the evaluation problem for a binary multiplexer. A hyperscheme is a high-level parameterized algorithm for solving a certain class of problems. Setting parameter values and subsequent interpretation of the hyperscheme allows obtaining algorithm schemes adapted to specific conditions of their use.

**The practical significance** of obtained results is that the means of hyperschemes are implemented in the developed integrated toolkit of automated design and synthesis of programs. Based on algorithm schemes, the system generates programs in a target programming language. The advantage of the system is the possibility of describing algorithm schemes in a natural-linguistic form. An experiment was conducted consisting in execution of the generated program for the problem of evaluating a binary multiplexer on a distributed cloud platform, which demonstrated the possibility of the developed distributed system to perform evaluations on 64 cloud clients-executors and obtain an increase in 60–100% of the maximum capabilities of a single-processor local implementation.

**Prospects for further research** are to apply the algebra-algorithmic method and tools for the automated development of the parallel implementation of evolutionary algorithm evaluation procedure on a graphics processing unit.

### ACKNOWLEDGEMENTS

The work is supported by the state budget scientific research project of the Institute of Software Systems of the National Academy of Sciences of Ukraine “Development of methods, technologies and tools for automating parallel programming using methods of computational intelligence” (state registration number 0122U002282).

УДК 004.4'24, 004.896

## ПАРАМЕТРИЧНО-КЕРОВАНА ГЕНЕРАЦІЯ ПРОГРАМИ ОЦІНКИ ДЛЯ АЛГОРИТМУ НЕЙРОЕВОЛЮЦІЇ НА ПРИКЛАДІ ДВІЙКОВОГО МУЛЬТИПЛЕКСОРА

**Дорошенко А. Ю.** – д-р фіз.-мат. наук, професор, завідувач відділу теорії комп'ютерних обчислень Інституту програмних систем НАН України, Київ, Україна.

**Ашур І. З.** – аспірант кафедри інформаційних систем та технологій Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна.

© Doroshenko A. Yu., Achour I. Z., Yatsenko O. A., 2023  
DOI 10.15588/1607-3274-2023-1-8

### REFERENCES

1. Doroshenko A. Formal and adaptive methods for automation of parallel programs construction: emerging research and opportunities / A. Doroshenko, O. Yatsenko. – Hershey : IGI Global, 2021. – 279 p. DOI: 10.4018/978-1-5225-9384-3
2. Designing neural networks through neuroevolution / [K. O. Stanley, J. Clune, J. Lehman et al.] // Nature Machine Intelligence. – 2019. – Vol. 1. – P. 24–35. DOI: 10.1038/S42256-018-0006-Z
3. SharpNEAT – Evolution of Neural Networks [Electronic resource]. – Access mode: <https://github.com/colgreen/sharpneat>
4. BinaryElevenMultiplexerEvaluator [Electronic resource]. – Access mode: <https://github.com/colgreen/sharpneat/blob/master/src/SharpNeatDomains/BinaryElevenMultiplexer/BinaryElevenMultiplexerEvaluator.cs>
5. Achour I. Z. Distributed implementation of neuroevolution of augmenting topologies method / I. Z. Achour, A. Yu. Doroshenko // Problems in Programming. – 2021. – № 3. – P. 3–15. DOI: 10.15407/pp2021.03.003
6. Yushchenko K. L. Algebraic-grammatical specifications and synthesis of structured program schemas / K. L. Yushchenko, G. O. Tseytlin, A. V. Galushka // Cybernetics and Systems Analysis. – 1989. – Vol. 25, № 6. – P. 713–727. DOI: 10.1007/BF01069770
7. Jiang D. Generation of C++ Code from Isabelle/HOL Specification / D. Jiang, B. Xu // International Journal of Software Engineering and Knowledge Engineering. – 2022. – Vol. 32, № 07. – P. 1043–1069. DOI: 10.1142/S0218194022500401
8. Moreira G. Fully-tested code generation from TLA+ specifications / G. Moreira, C. Vasconcellos, J. Knies // Systematic and Automated Software Testing : 7th Brazilian Symposium SAST'22, Uberlandia, 3–7 October 2022 : proceedings. – New York : ACM, 2022. – P. 19–28. DOI: 10.1145/3559744.3559747
9. Bonfanti S. Design and validation of a C++ code generator from abstract state machines specifications / S. Bonfanti, A. Gargantini, A. Mashkoor // Journal of Software: Evolution and Process. – 2020. – Vol. 32, № 2. – P. 1–27. DOI: 10.1002/smr.2205
10. Motwani M. Automatically generating precise oracles from structured natural language specifications / M. Motwani, Y. Brun // Software Engineering : 41st IEEE/ACM International Conference ICSE'2019, Montreal, 25–31 May 2019 : proceedings. – Los Alamitos : IEEE, 2019. – P. 188–199. DOI: 10.1109/ICSE.2019.00035
11. Kenneth O. S. A hypercube-based encoding for evolving large-scale neural networks / O. S. Kenneth, D. Ambrosio, J. Gauci // Artificial Life. – 2009. – Vol. 15, № 2. – P. 185–212. DOI: 10.1162/artl.2009.15.2.15202

Received 15.11.2022.  
Accepted 01.12.2022.

**Яценко О. А.** – канд. фіз.-мат. наук, старший науковий співробітник відділу теорії комп'ютерних обчислень Інституту програмних систем НАН України, Київ, Україна.

#### АНОТАЦІЯ

**Актуальність.** Розглянуто задачу автоматизованої розробки програм оцінки для алгоритмів нейроеволюції наростаючої топології. Еволюційні алгоритми застосовують механізми мутації, рекомбінації та селекції для пошуку нейронних мереж з поведінкою, яка задовольняє умовам певної формально визначеної задачі. Прикладом такої задачі є знаходження нейронної мережі, що реалізує певну цифрову логіку.

**Мета роботи** – автоматизоване проектування та генерація програми оцінки для задачі нейроеволюції на прикладі двійкового мультиплектора.

**Метод.** Методи та інструментальні засоби алгебри алгоритмів Глушкова та алгебри гіперсхем застосовано для параметрично-керованої генерації програми оцінки алгоритму нейроеволюції для бінарного мультиплектора. Алгебра Глушкова покладена в основу алгоритмічної мови, призначеної для багаторівневого структурного проектування та документування послідовних і паралельних алгоритмів та програм у формі, наближеній до природної мови. Гіперсхеми є параметризованими високорівневими специфікаціями, призначеними для вирішення певного класу задач. Задавання значень параметрів і подальша інтерпретація гіперсхем дозволяє отримати алгоритми, адаптовані до конкретних умов їх використання.

**Результати.** Засоби гіперсхем реалізовано в розробленому інтегрованому інструментарії автоматизованого проектування та синтезу програм. На основі схем алгоритмів система генерує програми цільовою мовою програмування. Перевагою інструментарію є можливість опису схем алгоритмів у природно-лінгвістичній формі. Проведено експеримент з виконання згерованої програми для задачі оцінки двійкового мультиплектора на розподіленій хмарній платформі. Згадана програма входить до складу SharpNEAT – системи з відкритим кодом, що реалізує алгоритм генетичної нейроеволюції NEAT для платформи .NET. Паралельна розподілена реалізація SharpNEAT була запропонована в попередній роботі авторів.

**Висновки.** Результати проведених експериментів продемонстрували можливість розробленої розподіленої системи виконувати оцінювання на 64 хмарних клієнтах-виконувачах та отримувати приріст у 60–100 % від максимальних можливостей однопроцесорної локальної реалізації.

**КЛЮЧОВІ СЛОВА:** алгебра алгоритмів, автоматизоване проектування програм, хмарні обчислення, гіперсхема, нейроеволюція, нейронна мережа, паралельне програмування.

#### ЛІТЕРАТУРА

1. Doroshenko A. Formal and adaptive methods for automation of parallel programs construction: emerging research and opportunities / A. Doroshenko, O. Yatsenko. – Hershey : IGI Global, 2021. – 279 p. DOI: 10.4018/978-1-5225-9384-3
2. Designing neural networks through neuroevolution / [K. O. Stanley, J. Clune, J. Lehman et al.] // Nature Machine Intelligence. – 2019. – Vol. 1. – P. 24–35. DOI: 10.1038/S42256-018-0006-Z
3. SharpNEAT – Evolution of Neural Networks [Електронний ресурс]. – Режим доступу: <https://github.com/colgreen/sharpneat>
4. BinaryElevenMultiplexerEvaluator [Електронний ресурс]. – Режим доступу: <https://github.com/colgreen/sharpneat/blob/master/src/SharpNeatDomains/BinaryElevenMultiplexer/BinaryElevenMultiplexerEvaluator.cs>
5. Ашур І. З. Розподілена реалізація методу нейроеволюції наростаючої топології / І. З. Ашур, А. Ю. Дорошенко // Проблеми програмування. – 2021. – № 3. – С. 3–15. DOI: 10.15407/pp2021.03.003
6. Yushchenko K. L. Algebraic-grammatical specifications and synthesis of structured program schemas / K. L. Yushchenko, G. O. Tseytlin, A. V. Galushka // Cybernetics and Systems Analysis. – 1989. – Vol. 25, № 6. – P. 713–727. DOI: 10.1007/BF01069770
7. Jiang D. Generation of C++ Code from Isabelle/HOL Specification / D. Jiang, B. Xu // International Journal of Software Engineering and Knowledge Engineering. – 2022. – Vol. 32, № 07. – P. 1043–1069. DOI: 10.1142/S0218194022500401
8. Moreira G. Fully-tested code generation from TLA+ specifications / G. Moreira, C. Vasconcelos, J. Knies // Systematic and Automated Software Testing : 7th Brazilian Symposium SAST'22, Uberlandia, 3–7 October 2022 : proceedings. – New York : ACM, 2022. – P. 19–28. DOI: 10.1145/3559744.3559747
9. Bonfanti S. Design and validation of a C++ code generator from abstract state machines specifications / S. Bonfanti, A. Gargantini, A. Mashkoor // Journal of Software: Evolution and Process. – 2020. – Vol. 32, № 2. – P. 1–27. DOI: 10.1002/smr.2205
10. Motwani M. Automatically generating precise oracles from structured natural language specifications / M. Motwani, Y. Brun // Software Engineering : 41st IEEE/ACM International Conference ICSE'2019, Montreal, 25–31 May 2019 : proceedings. – Los Alamitos : IEEE, 2019. – P. 188–199. DOI: 10.1109/ICSE.2019.00035
11. Kenneth O. S. A hypercube-based encoding for evolving large-scale neural networks / O. S. Kenneth, D. Ambrosio, J. Gauci // Artificial Life. – 2009. – Vol. 15, № 2. – P. 185–212. DOI: 10.1162/artl.2009.15.2.15202



## THE CURVE ARC, AS A STRUCTURE ELEMENT OF AN OBJECT CONTOUR IN THE IMAGE TO BE RECOGNIZED

**Kalmykov V. G.** – PhD, Senior Researcher of the Institute of Mathematical Machines and Systems Problems, Kyiv Ukraine.

**Sharypanov A. V.** – PhD, Senior Researcher of the Institute of Mathematical Machines and Systems Problems, Kyiv Ukraine.

**Vishnevskiy V. V.** – PhD, Leading Researcher of the Institute of Mathematical Machines and Systems Problems, Kyiv Ukraine.

### ABSTRACT

**Context.** The proposed article relates to the field of visual information processing in a computer environment, more precisely to the determination the parameters of the interest object in the image, in particular, the contour of the interest object In most cases, the contour of an object is a simply connected sequence of curve arcs.

**Objective.** The purpose and subject of the study is to find and to propose such a definition of the digital curve arc, as the most important element of the object contour in the recognizable image, which does not contradict modern neurophysiological conceptions about visual perception, and to recognize the object contour as a sequence of the digital curve arcs.

As a **Method**, the representation of the image in the form of a structural model is used, one of the structural elements of which is the contour of the object, consisting of digital curve arcs. Also, the image is considered as a cellular complex which corresponds to modern ideas about human visual perception.

**Results.** The new definition for arc of a digital curve as a sequence of digital straight segments is proposed, which does not contradict to modern concepts of neurophysiology. In contrast to the known definitions of a curve arc, the proposed definition of a digital curve arc makes it possible to determine the start and end points of the arc. According to the description of the contour of an object as a simply connected closed sequence of line segments, it is proposed to construct a description of the contour as a sequence of arcs of digital curves.

**Conclusions.** The use of the proposed definition of the digital curve arc in image processing makes it possible to recognize the contour of an object in an image and present it in a form close to visual perception. For best results, the use of variable resolution in image processing algorithms is recommended.

**KEYWORDS:** image, contour, curve arc, straight line segment, cellular complex, neurons, receptive field.

### NOMENCLATURE

$AB$  is the curve arc, corresponds to  $x = \varphi(\tau)$ ,  $y = \phi(\tau)$  at  $l \leq \tau \leq u$ ;

$a$  is an example of 0-cell;

$b$  is an example of 0-cell;

$c$  is an example of 0-cell;

$C$  is an abstract cell complex;

$C_m$  is an  $m$ -dimensional Cartesian complex;

$d$  is the pixel size;

$dim$  is a dimension function;

$E$  is the set of abstract elements (cells);

$e$  is an example of 0-cell;

$e'$  is formal cell example;

$e''$  is formal cell example;

$f$  is an example of 0-cell;

$F$  is a binary relation;

$g$  is an example of 0-cell;

$I$  is the set of non-negative integers

$h$  is the height of the arc segment;

$h_s$  is the height of the  $s$ -th arc segment;

$l$  is lower bound of the parameter  $\tau$  definition area;

$m$  is a dimension;

$n$  is the cell number in the cell sequence;

$N$  is the cell quantity in the cell sequence;

$p$  is an example of 0-cell;

$P$  is a set of points;

$r$  is a brightness function;

$s$  is the straight segment number in the sequence;

$S$  is the straight segment quantity in the sequence;

$t$  is the cell number in the cell sequence of the straight line segment;

$T_b T_e$  is a straight line segment;

$T_b$  is a begin point  $x_b, y_b$  of the straight line segment  $T_b T_e$ ;

$T_e$  is an end point  $x_e, y_e$  of the straight line segment  $T_b T_e$ ;

$u$  is upper bound of the parameter  $\tau$  definition area;

$x_n$  is the cell abscissa with number  $n$  in the sequence;

$y_n$  is the cell ordinate with number  $n$  in the sequence;

$x_s$  is the abscissa of the boundary common point (0-cell) of adjacent line segments, number  $s$  in the sequence;

$y_s$  is the ordinate of the boundary common point (0-cell) of adjacent line segments, number  $s$  in the sequence;

$x_s$  is the boundary point (0-cell) number  $s$  abscissa of line segment belonging to curve arc;

$y_s$  is the boundary point (0-cell) number  $s$  ordinate of line segment belonging to curve arc, in the sequence;

$x_t$  is the cell abscissa with number  $t$  in the sequence of the straight line segment;

$y_t$  is the cell ordinate with number  $t$  in the sequence of the straight line segment;

$\varphi$  is continuous parametrically defined function;

$\psi$  is continuous parametrically defined function;  
 $\tau$  is the parameter on the segment  $[L, u]$ .

## INTRODUCTION

Many tasks of image analysis and processing consist in detecting an object of interest and determining its parameters. Typically, visual information and, in particular, a grayscale image is represented as a set of pixels, densely, without gaps, filling the field of the image. The boundaries (contour) of the detected object are often not defined as the result of image processing, as, for example, for statistical [1, 2, 3, 4] or neural network [5, 6, 7] approaches to image processing.

At the same time, other approaches are known for solving image recognition tasks, in particular, structural recognition [8], which involve the representation of an image in the form of a hierarchical structure.

**The object of study** is the arc of the curve as an structural element of the image while its processing.

**The subject of study** is the development of the curve definition arc, suitable for the analysis and processing of an image in its discrete representation. The well-known mathematical definitions of the curve arc are not very suitable for use in discrete image processing, since they define abstract curves in a continuous space.

**The purpose of the work** is to develop a definition of the curve arc, which is suitable for finding the arc of the curve in the image in the process of recognizing the contour of an object.

## 1 PROBLEM STATEMENT

Let  $\{x_n, y_n\}$ ,  $1 \leq n \leq N$ , be a sequence of points (0-cells) describing the contour of an object. Let the contour of the object be approximated by line segments and  $\{x_s, y_s\}$ ,  $1 \leq s \leq S$ , be a sequence representing boundary points (0-cells) of adjacent line segments. That is, each adjacent two line segments are represented in the sequence by pairs of points (0-cells):  $(x_{s-1}, y_{s-1}; x_s, y_s)$  and  $(x_s, y_s; x_{s+1}, y_{s+1})$ , respectively.

The problem: whether there are in a sequence of line segments such that they form an arc of a curve can be reduced to checking the condition of belonging or not belonging to an curve arc of two adjacent line segments, followed by checking all pairs of line segments in the sequence.

Then it is necessary and sufficient to find the condition of belonging or not belonging to the arc of the curve of two adjacent line segments in the sequence and apply it to all pairs of adjacent segments in the sequence.

## 2 REVIEW OF THE LITERATURE

Consider the halftone image structure. The first hierarchical structure of an image is discussed in [8]. A more general structural image model, compared to [8], is presented here. In most cases, a grayscale image can be considered as a realization of an unknown brightness function  $r = f(x, y)$  depending on two spatial variables  $x, y$ . This function defines a piecewise smooth surface. Objects in

the field of view are regular pieces of a piecewise-smooth surface. The projections of the surface pieces contours onto the image plane coincide with the definition domains boundaries of the unknown functions that define the surface pieces, and are the objects contours in the field of view. The objects forming the background are the objects in the field of view, the contours of which partially or completely coincide with the boundaries of the image. Possible objects of interest may include objects in the field of view that do not have common boundaries with the boundaries of the image. Visual information should be presented taking into account the physiological characteristics of visual perception so that optimal results can be achieved with its automatic processing. In particular, one of the most important and natural features of human visual perception is its ability to segment the visual field into objects that differ from the background in brightness, color, texture. The main characteristic of any object is its shape, determined by the contour – the boundary between the object and the background. The contour, in turn, is perceived by a person as a sequence of straight line segments and curve arcs. The shape of halftone objects is also determined by the brightness function based on the color, texture within each of the objects.

These features of human visual perception are reflected in the structural model of the image. The structural model makes it possible to represent arbitrary images uniformly in form, invariant with respect to affine transformations – position in the field of view, scale, rotation. The problem of reduction to a structural model of arbitrary images given in a raster form, distorted by noise, in the general case, has not yet been solved. However, the transformation of images into a structural model can significantly increase the speed and quality of visual information processing in some rather numerous cases. The basis for the structural analysis of a halftone image is a model that determines its structural elements (Fig. 1). In particular, the objects of interest and the background of the image are such structural elements according to the known ideas about the mechanisms of visual perception. Objects, in turn, are defined by bounding contours and a three-dimensional brightness function within the object.

Contours are closed sequences formed by segments of straight lines and arcs of curved lines. Representation of grayscale images in the form of such or similar model is invariant to affine transformations of objects.

Consider the image as a cellular complex. There are many problems in image analysis that cannot be solved based on classical Euclidean geometry. The reason is that in classical geometry the assumption of space continuity is used. That is, each point in space contains in its neighborhood an infinite number of points, no matter how small this neighborhood is. According to the topological foundations of classical geometry, even the smallest neighborhood of each point contains an infinite number of other points. Thus, classical geometry has no means for processing discrete images, because discrete image is presented as the set of isolated points, sufficiently small

neighborhoods of which do not contain points at all, except for the isolated point itself. But then the discrete image can be described in classical geometry very approximately, with an accuracy of several distinctly small spatial elements.

It was proposed to use the mathematical apparatus of abstract cellular complexes to describe the image [9, 10]. An abstract cellular complex  $C = (E, F, \dim)$  is understood as a set  $E$  of abstract elements (cells) that are in an antisymmetric, irreflexive and transitive binary relation  $F \subset E \times E$ , which is called the limit relation, and with a dimension function  $\dim: E \rightarrow I$  with  $E$  on the set  $I$  of non-negative integers, such as  $\dim(e') < \dim(e'')$  for all pairs  $(e', e'') \in F$ . If  $(e', e'') \in F$  then one usually writes  $e' < e''$  or says that cells  $e'$  limit cells  $e''$ . Such cells are called incident to each other.

Fig. 2a shows an arbitrary halftone image, which can be described as an abstract two-dimensional cellular complex. First of all, let's note the 0-dimensional cells:  $a, b, c, d, e, f, g$ . 0-dimensional cells correspond to points in the two-dimensional Euclidean space. The attributes of each point are coordinates. Lines correspond to 1-dimensional cells in Euclidean space:  $ab, bc, cd, de, ef, fg, ga$ . These lines in Euclidean space correspond to an interval – an open set of points, the closure of which is the

points corresponding to 0-dimensional cells. For example, a 1-dimensional cell  $ab$  is bounded by 0-dimensional cells (corresponding to points in the two-dimensional Euclidean space)  $a, b$ . One-dimensional cells have no thickness. The attributes of each line are the coordinates of its boundary points. The pieces of the plane correspond to 2-dimensional cells  $abg, bcfg, cdf, def$  in the Euclidean space. The sets of points that form pieces of the plane are open. 2-dimensional cells are bounded by the corresponding 1-dimensional and 0-dimensional cells. The sets of points corresponding to 1-dimensional and 0-dimensional cells, bounding the 2-dimensional cells, are the closures of the point sets of these 2-dimensional cells. For example, a 2-dimensional cell  $abg$  is bounded by 1-dimensional cells  $ab, bg, ga$  and 0-dimensional cells  $a, b, g$ . The attributes of each piece of plane are the brightness value of its points, the coordinates of the bounding points and lines.

The discretized image can be represented as a Cartesian two-dimensional cell complex [10]. Each coordinate axes can be considered as a sequence of 0-cells and alternating 1-cells in the space where the image is presented (Fig. 2b). 0-cells are assigned to the points of intersection

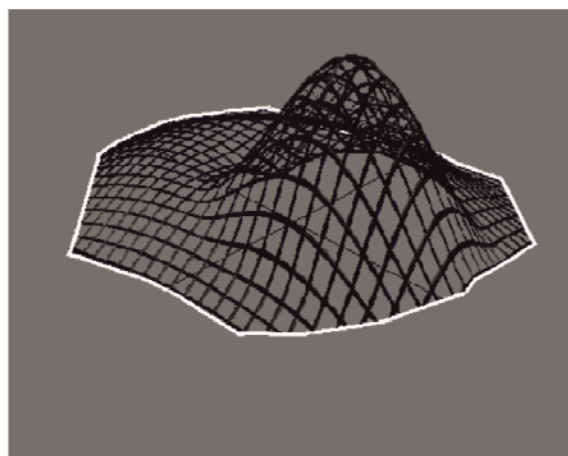
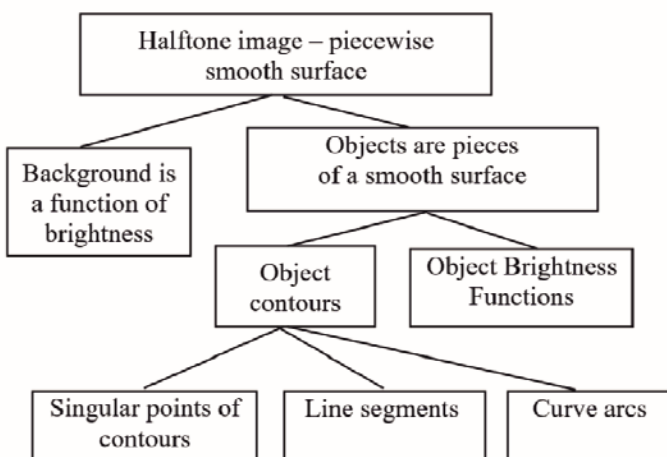


Figure 1 – The structural model of halftone image

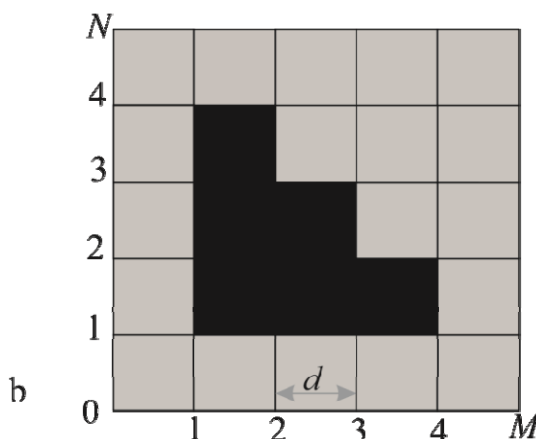
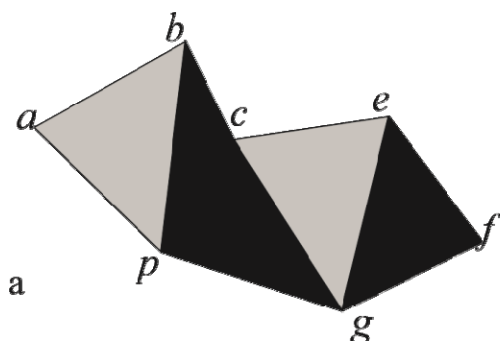


Figure 2 – a – image as a cell complex; b – image as a Cartesian cell complex;

of the grid lines with the axis, 1-cells are assigned to the segments between adjacent points of intersection of the grid lines with the axis. The set of cells of a two-dimensional Cartesian complex is the Cartesian product of sets of axes cells. This means that a cell of the  $m$ -dimensional Cartesian complex  $C_m$  is an  $m$ -tuple of axis cells. The bound ratio in  $C_m$  is derived from the boundary ratios of the axes. The dimension of the cell  $C_m$  is the sum of dimensions of its factors. 2-dimensional cells or grid cells correspond to pixels.

In the object in an image description the attributes of 0-cell are its coordinates, the attributes of 1-cell are its initial and final coordinates. The attributes of a 2-cell are the coordinates of its boundary 0-cell closest to the origin, as well as the brightness value of the corresponding image pixel. Each 2-cell corresponding to an image pixel is bounded by four 1-cells – cracks and four 0-cells – dots. Each 1-cell is bounded by two 0-cells. The most important feature of the object on an image is its boundary – contour. Contour of the object, when presented as a cellular complex, is a closed sequence of 0-cells and 1-cells. Fig. 2b shows the object of the halftone image and the contour of the object in the form of sequence of 0-cells and 1-cells. So, the contours are represented by lines without thickness. Thus, the description of a halftone image as a Cartesian cell complex contains sets of 0-cells, 1-cells, and 2-cells. These sets can be represented in computer memory as separate arrays.

Populations of neurons [11] have been found in the striate cortex, whose neuronal responses form the contours of figures in the visual field if their receptive fields correspond to the boundary segments of contrasting objects. The properties of these neurons' responses correspond to properties of 1-dimensional cells of cell complexes. These neurophysiological studies were carried out completely independently of mathematical work in the field of image description by methods of discrete topology.

The functioning of neurons comparable to 1-dimensional cells is described below. Most cortical cells (neurons) respond poorly to diffuse illumination and respond well to contrasting borders with a suitable orientation. For an arbitrary figure contrasting with the background, such a cell will respond if and only if a boundary segment with a certain orientation intersects its receptive field. "The same cells, whose receptive fields are located inside the boundaries of the figure, will not react in any way – they will continue to give a spontaneous impulse discharge regardless of the presence or absence of this figure. However, to excite a simple cell, it is not enough that the boundary section corresponds to the optimal orientation – the contour must also almost exactly hit the edge of the inhibitory and excitatory zones of the receptive field, because it is necessary for the answer that the light falls on the excitatory zone, but does not spread to the inhibitory one. If you the section of the contour was shifted even slightly without changing its orientation, the stimulation of this cell will turn out to be insufficient, and now another population of simple cells will begin to ex-

cite". It is known that each point of the visual field corresponds to a set of neurons (cells) with different orientational selectivity. The cells whose receptive fields and orientation coincide with the contour of the figure will answer to stimulus. The set of responses of such cells completely describes the contour of the figure. The following is also noteworthy: the output signals of cells whose receptive fields coincide with the contour of the figure and are oriented accordingly correspond to sections of the contour, one-dimensional segments without line width. That is, a material object – a neuron and its receptive field, corresponding to a part of the image, put in correspondence with an intangible object of a straight line segment – an element of the contour. Each of these contour elements can be considered as a 1-dimensional cell of the cellular complex. Here is another quote from [11]: "Another type of cells is found in the striate cortex. Typically, simple and complex cells are characterized by spatial summation – the longer the stimulus line, the better the response. However, the response only intensifies until the length of the line reaches the size of the receptive field: further lengthening of the line does not lead to a more vigorous response. In contrast, in cells that respond to line ends (end stopped cells), lengthening the line to a certain limit continues to improve the response, and if the line goes beyond this limit (in one or both directions), then the response weakens. Some cells, which we call "completely end stopped cells", do not respond at all to the presentation of a stimulus in the form of a long line. The zone from which a cell response can be elicited is called the activation zone (or excitatory zone), and the zones located at one or both ends are called inhibition zones (inhibitory zones). Thus, the entire receptive field of such a cell consists of an excitatory zone and an inhibitory zone (or zones) at the edges. A stimulus of optimal orientation, activating a cell from the excitatory zone, causes maximum inhibition outside this zone (on one or both sides)." From the standpoint of representing an image as a cellular complex, the above example can be considered as an experimental confirmation of the implementation of 0-dimensional cells in the visual system of a living organism.

Consider the mathematical definitions of the curve arc. Usually the objects contours are presented for further analysis as a closed sequence of curves arcs. The concept of a curve arc (meaning a continuous curve) is used in various fields of science and technology, in particular, recently in the processing of visual information. The concept of a continuous curve is one of the concepts that seems intuitively simple, but is actually very difficult to define. The greatest mathematicians defined the continuous curve in various ways at different periods in the development of this field of knowledge. Each new definition proceeded from the needs of human practical activity and the mathematical knowledge level of the corresponding era. The most modern definitions closely related to set theory are as follows [12]:

Definition (according to Jordan). A plane curve is a set of points in a plane whose coordinates are determined by two equations:

$$\begin{aligned} x &= \varphi(\tau), \\ y &= \psi(\tau), \end{aligned}$$

where  $\tau$  is defined on some segment  $[l, u]$ . The choice of the segment values does not violate the generality of the definition.

Definition (according to Cantor) [13]. A curve in a plane is any connected, compact set  $P$  of points in the plane that does not contain any interior point.

Definition (according to Urysohn) [14]. A curve is a one-dimensional connected and at the same time compact set.

The following difficulties occur when using the above definitions while processing the visual information.

1. It is assumed in the given definitions of continuous curves that the exact boundaries of the segment or data set to be examined for suitability or non-suitability with these definitions are known. But this part of the total data set that corresponds to one or another curve arc is not known in advance, when processing signals and visual information. This part of the data set can only be formed during processing. For example, when processing an image, it is not known which part of pixels belongs to the object of interest, or to its boundaries. Often it is obtained as the result of image processing.

2. It is essential to consider the difference between curve arcs and line segments when processing visual information. This difference is not provided in the above definitions.

3. The above definitions are found and fulfilled for continuous curves in a continuous Euclidean space. But in the modern view, visual information, signals, as well as visual perception, are inherently discrete. That is, we are talking about a discrete space, where each point is isolated.

4. Traditionally, the points of a discrete curve arc in a real image are represented by pixels that have real dimensions, while the points of mathematical curves refer to infinitesimal values.

It should be recognized that another definition is required for the visual information and signals processing, taking into account the above considerations, while fully recognizing the great scientific and practical significance of the known definitions of a continuous curve.

### 3 MATERIALS AND METHODS

In general, the task of object contour recognition as finding simply connected closed sequence of 0-dimensional and 1-dimensional cells includes processing of grayscale image using variable resolution. The presentation of this problem and its solution are beyond the scope of this work. At the same time, the solution of this problem for a binary image is given in [15]. In the same way, the construction of an object contour for a large number of halftone images is reduced to the mentioned

problem if it is possible to binarize a halftone image using a specially selected threshold. We will assume that for the object of interest in a grayscale image, a contour is constructed as a simply connected closed sequence of 0-dimensional and 1-dimensional cells. But 0-dimensional and 1-dimensional cells are elements of the contour, commensurate with the pixel, which in turn is not structural semantic, meaningful element of the contour.

As already mentioned, in accordance with the structural model of a halftone image, the contour is a simply connected closed sequence of structural elements – straight line segments and arcs of curve, which are formed from parts of a sequence of 0-dimensional and 1-dimensional cells. The task is to represent the object contour as a closed simply connected sequence of line segments and arcs of curve using the mentioned closed simply connected sequence of 0-dimensional and 1-dimensional cells as input data.

First of all, the simply connected closed sequence of line segments must be calculated for the sequence of 0-dimensional and 1-dimensional cells. Each segment corresponds to a certain simply connected part of the original sequence. The definition of line segments must be performed by 0-dimensional cells sequentially, starting from the first 0-dimensional cell of the sequence. Let  $\{x_n, y_n\}$ ,  $n=\overline{1, N}$  be a sequence of coordinates of 0-dimensional cells.

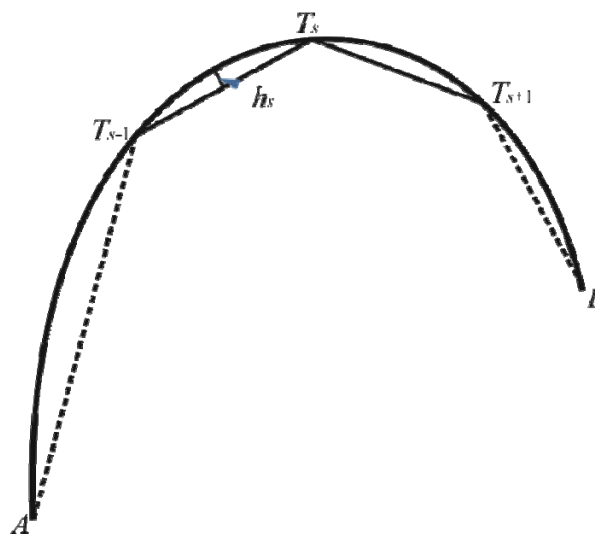


Figure 3 – The curve arc with an inscribe polyline

The possibility of representing the next part of a sequence of 0-dimensional cells  $\{(x_b, y_b), \dots, (x_t, y_t), \dots, (x_e, y_e)\}$ , between the begin point  $x_b, y_b$  and end point  $x_e, y_e$ , as a segment of a straight line with the begin point  $T_b=(x_b, y_b)$  and end point  $T_e=(x_e, y_e)$ , is determined by the condition [16] that

$$\max_{\forall (b \leq t \leq e)} \text{dist}(T_b T_e, (x_t, y_t)) \leq d/2,$$

here  $dist(T_b T_e, (x_t, y_t))$  – distance of the cell  $(x_t, y_t)$  to a line segment  $T_b T_e$ . That is, the distance from any 0-dimensional cell belonging to segment  $T_b T_e$  should not exceed half the length of a 1-dimensional cell. Consistent application of the above condition allows us to represent the entire sequence as a simply connected sequence of line segments or  $\{(x_1, y_1), \dots, (x_s, y_s), \dots, (x_S, y_S)\}$  – a sequence of boundary common 0-cells (points) of adjacent line segments. Thus, the contour of the object is presented as a simply connected, closed sequence of digital straight segments.

The definition of the arc of a digital curve proposed below makes it possible to establish or reject the fact that a sequence of digital straight segments is received due to that some arc of the curve has been discretized. We will assume, that the arcs of the curve used in graphic images represent segments of smooth functions and correspond to Jordan curves. The arcs of an arbitrary curve [9], are given by the equations  $x = \varphi(\tau)$ ,  $y = \phi(\tau)$ , without multiple points or simple arcs, that is, such that for any two different values  $\tau'$  and  $\tau''$  the corresponding points on the plane  $[\varphi(\tau'), \phi(\tau')]$  and  $[\varphi(\tau''), \phi(\tau'')]$  are different. Let  $x = \varphi(\tau)$ ,  $y = \phi(\tau)$ , where the parameter  $\tau$  defined on the segment  $[l, u]$ . As  $\tau$  increases from  $l$  to  $u$ , the point with coordinates  $x, y$  describes the arc  $AB$  (Fig. 4). Consider a partition of the segment  $[l, u]$  by division points

$$l = t_0 < \dots < t_{s-1} < t_s < t_{s+1} < \dots < t_S = u,$$

and let these points of division correspond to the points of the curve  $A, \dots, T_{s-1}, T_s, T_{s+1}, \dots, B$ .

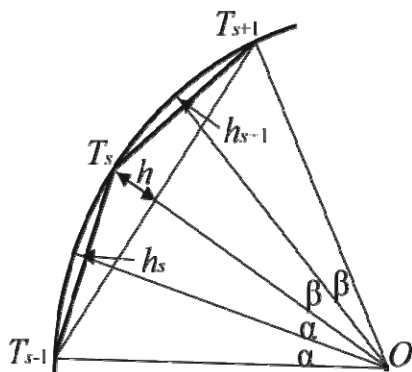


Figure 4 – The correspondence checking of straight line segments pairs to the definition of an curve arc

A polyline, inscribed in the arc  $AB$ , will be constructed if we connect successively point  $A$  with point  $T_1$ , point  $T_1$  with... point  $T_{s-1}$ , point  $T_{s-1}$  with point  $T_s$ , point  $T_s$  with point  $T_{s+1}$ , point  $T_{s+1}$  with ... point  $B$  by segments of straight lines. The figure bounded by the segment of the polyline  $T_s, T_{s+1}$  and the corresponding arc link  $\cap T_s, T_{s+1}$  will be called the segment of the arc  $T_s, T_{s+1}$ , and the maximum length of the line between the segment  $T_s, T_{s+1}$  and  $\cap T_s, T_{s+1}$ , perpendicular to the segment  $T_s, T_{s+1}$  is the height of the arc segment  $h_s$ . Let be

$$\beta = \max_{s=0,1,\dots,S-1} l(T_s, T_{s+1}).$$

If  $\beta$  tends to zero with a corresponding increase in  $s$ , then the length of any of the links of the inscribed polyline will tend to zero, as well as the height of each segment of the arc, due to the continuity of the functions  $\varphi(t), \phi(t)$ .

While an arc and an inscribed polyline represent in a discrete space of discreteness  $d$ , segments of the inscribed polyline are displayed as line segments. Since the coordinate values take integer multiples of  $d$  in a discrete space, then objects smaller than half the discreteness the heights of the segments, in particular, will not be displayed in this space, their lengths will become equal to zero, starting from the moment when  $h_s < d/2$ . So, the discrete mappings of the parts of the arc will coincide with the corresponding links of the inscribed polyline – segments of digital lines for  $h_s < d/2$ . Thus, the contour, which consists of straight segments and arcs of arbitrary curves, after discretization is defined as a sequence of digital straight segments. Sequences of digital straight segments that correspond to arcs of curves can be considered as polylines inscribed in these arcs of curves. Such inscribed polylines will be called arcs of digital curves. The contour can include both individual segments of straight lines, and sequences of such segments – polylines that are not arcs of digital curves.

Consider pairs of adjacent segments of digital lines in sequence. In general, many curves can be drawn through three points defined by a pair of line segments. Nevertheless, as already noted, the lengths of the arcs segments heights that correspond to the segments of the inscribed polyline should not exceed the value of the space discreteness  $d/2$ . Thus, in order to consider pairs of segments of digital lines  $T_{s-1}T_s, T_sT_{s+1}$  as part of an arc of a digital curve, it is necessary to establish the existence of a curve that passes through the points  $T_{s-1}, T_s, T_{s+1}$ , such that the condition is satisfied:  $(h_s < d/2) \& (h_{s+1} < d/2)$  (Fig. 4).

The curvature of a plane curve is usually identified with the curvature of a contacting circle [9]. The contacting circle of a plane curve at the point  $T_s$  is the limiting position of the circle passing through two neighboring points  $T_{s-1}$  and  $T_{s+1}$  as  $T_{s-1}$  and  $T_{s+1}$  tend to  $T_s$ . We can formulate the following definition, based on the above considerations.

**Under the arc of a digital curve in a two-dimensional discrete space of discreteness  $d$  we mean such a sequence of straight line segments that through the three end points of each pair of adjacent segments it is possible to draw such a circle that the heights of the circle segments do not exceed  $d/2$ .**

This definition is valid to the extent that it is legitimate to identify a segment of an arc of an arbitrary curve that corresponds to a pair of neighboring segments with an arc of a contacting circle.

Having constructed a circle in accordance with the definition of the digital curve arc for the points  $T_{s-1}, T_s, T_{s+1}$ , let us estimate the distance of the common

point  $T_s$  of the pair of segments  $(T_{s-1}, T_s), (T_s, T_{s+1})$  to the segment  $T_{s-1}, T_{s+1}$ , that is, the height  $h$  (Fig. 4). The lengths of each segment in this pair cannot differ significantly, since this would contradict the smoothness condition – that is,  $l(T_{s-1}, T_s) \sim l(T_s, T_{s+1})$ . As already noted, the maximum distance between the points of the arc lines and the corresponding segment of the digital straight line is  $h_s = h_{s+1} = d/2$ . At the same time

$$h_s = OT_{s-1} - OT_{s-1} \times \cos \alpha = r - r \cos \alpha = r(1 - \cos \alpha). \\ h = OT_{s-1} - OT_{s-1} \times \cos 2\alpha = r - r \cos 2\alpha = \\ = r(1 - \cos 2\alpha) = 2r(1 - \cos^2 \alpha).$$

$h/h_{s-1} = 2(1 + \cos \alpha)$ ; or  $h = 2(1 + \cos \alpha) \times h_{s-1}$ . If  $h_{s-1} \approx d/2$  and  $\alpha \leq 10^\circ$ , that is  $\cos \alpha \approx 1$ , then the height of the triangle  $(T_{s-1}, T_s, T_{s+1})$   $h \approx 2d$ . Using the value  $h_s$  instead of  $h_{s-1}$  will not affect the result, since both  $\beta \leq 10^\circ$  and  $\cos \beta \approx 1$ .

This means that in order to be related to the digital arc of curve for the considered pair of segments, it is necessary that the value of the maximum deviation of  $h$  does not exceed  $2d$ . The minimum deviation is  $h > d/2$ , since at a smaller deviation the directions of the segments  $T_{s-1}T_s$  and  $T_sT_{s+1}$  are indistinguishable, and a pair of segments of different directions turns into one straight segment. If  $h > 2d$ , then the segments under consideration are segments of a polyline. Thus, taking into account the above considerations, the sequence of common points of adjacent segments takes the form:  $\{(x_1, y_1), \dots, (\underline{x_{s-1}}, \underline{y_{s-1}}), (\underline{x_s}, \underline{y_s}), (\underline{x_{s+1}}, \underline{y_{s+1}}), \dots, (x_S, y_S)\}$ , where the points belonging to the digital curve arc are underlined.

#### 4 EXPERIMENTS

Experimental verification of the proposed method consists in representing the contour of the interest object as a sequence of digital curve arcs and segments of digital straight lines. Moreover, the contour elements sequence of the object must be the same for various affine transformations – the rotation of the interest object, changing the position in the image field. For comparison, the representation the contour of the interest object as a sequence of digital curve arcs and segments of digital straight lines was performed using a well-known tool – the graphical editor Corel Draw.

For the experiment, binary images of object contours that were not distorted by noise were used, since noise filtering, as well as recognition of the object contours in a grayscale image, are separate tasks that must be solved by appropriate means. Separate works will be devoted to solving these problems.

An example of the image used in the experiment is shown in Fig. 5. Objects in the image are the identical sectors of the ellipse, differing in space position and angle of rotation. The result of processing each object is its contour, represented by a closed sequence of line segments in the form of boundary points (0-cells) of adjacent line segments. Some parts of the line segments sequence are defined as arcs of curves with indication of their boundary points (0-cells).



Figure 5 – The example of the image for the experiment to determine the contour as a sequence of curve arcs and line segments

The experimental program was executed in the Visual C environment. The required RAM is no more than 512 MB.

The main blocks of the program:

1. Represent the image as a cell complex and determine the sequences of boundary 0-cells and 1-cells for each object.

2. Define the contour of each object as a sequence of line segments in the form of a sequence of boundary points of neighboring segments.

3. Determine for the contour of each object the parts of the line segments sequence that form the arcs of the curves, indicating the boundary points of the arcs of the curves. The use of curve arcs as the structural elements of image contours description would approach its description to intuitional, natural representation of images by a man, substantially would shorten the expenses of memory for storage of image and image processing time. As an example we will consider description of contours of binary images which are got with the use of tools of widespread graphics editor of Corel Draw. On fig.5a the contours of three identical objects are represented, not to be distorted by noises. Each of objects contains the arc of ellipse and differs from the other objects by spatial position and rotation angle. The boundary points which divide contours into the curve arcs and the straight segments are marked by the squares. Identical with each other arcs, to belong to different objects, are represented by sequences containing the different amount of different arcs of curves. Each of identical objects in the image is represented with the different elements. Such description of objects can not be directly used in the intelligence systems for interpretation of images, as supposes enough hard processing. The represented example shows existence of the problem even at the images not distorted by noises and actuality to solve the problem. A special program that implements the proposed method and algorithms has been developed.

The example of contour recognition as the sequences of digital straight segments and of digital curve arcs by the program is demonstrated on a Fig. 6b. The image is used from the Fig. 6a, but the offered algorithms are implemented in the program. Unlike the contours of Fig. 6a, got by means of the program of Corel Trace, the arcs of contours are represented without laying out by intermediate points on a few arcs regardless of different spatial positions and rotation angles for the each of objects

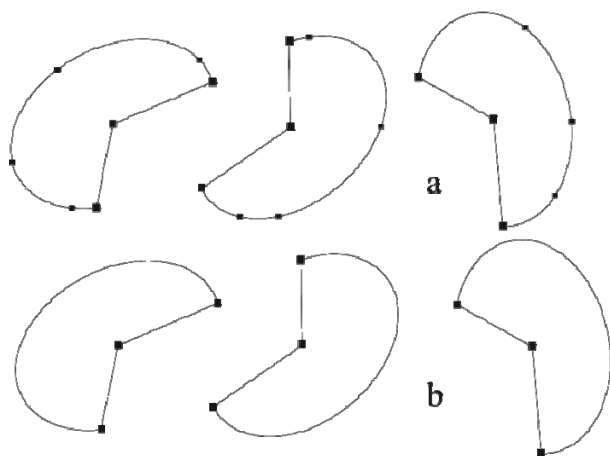


Figure 6 – a – Contour recognition in three identical, rotated in relation to each other objects by facilities of Corel Draw;  
b – Contour recognition by the program, using the proposed method and algorithms

## 5 RESULTS

The main result, apparently, should be considered the ability to determine the arcs of curves in sequences of 0-cells and 1-cells that form the contour of an object. The same contour configurations must correspond to the same sequences of curve arcs, regardless of affine transformations – shift, rotation angle, scale.

The significance of the result obtained is clear only when comparing the processing of the same object by known and proposed methods. The result of the experiment is shown in Fig. 6. The contours of objects from Fig. 5, recognized by means of the Corel Draw graphic editor are shown in fig. 6a. Each of objects contains the arc of ellipse and differs from the other objects by spatial position and rotation angle. The boundary points which divide contours into the curve arcs and the straight segments are marked by the squares. Identical with each other arcs, to belong to different objects, are represented by sequences containing the different amount of different arcs of curves. Each of identical objects in the image is represented with the different elements. The arc of each ellipse is represented by several unequal arcs. Such representation of processing results cannot be used in artificial intelligence tasks, in particular, in recognition tasks.

The proposed method is free from these defects. Recognition of the same objects by the developed experimental program is shown in Fig. 6b

## 6 DISCUSSION

The representation of image within the framework of the structural model, on the one hand, is natural for visual perception, on the other hand, it is fully consistent with the theory of cell complexes. The image object is a 2-dimensional cell. The contour of the object, its boundary, is, most often, a simply connected, closed sequence of 1-dimensional cells that form segments of straight lines and arcs of curve. The boundaries of line segments and arcs of curve, i.e. singular points correspond to 0-dimensional cells. Compared to recent imaging concepts, an important

advantage to using cell complexes in image processing is the following [16].

One of the simplest tasks of image processing is to encode the object contour of the binary image as a single-connected closed sequence of the object boundary elements in the image. The contour of the object is a closed curve that divides the image into two parts: the object itself and the other part of the image. Traditionally, for a binary discretized image, contour pixels bounding with background pixels or, conversely, background pixels bounding with object pixels are used as contour elements. In order to construct a single-connected closed sequence from boundary pixels that corresponds to the contour of the image object, the concept of connectivity – the pixels neighborhood – must be defined. The following ideas about the connectivity of pixels in two-dimensional discrete space are generally accepted:

1. Pixels are considered adjacent if they have a common side. In this case, each pixel has four adjacent pixels.
2. Pixels are considered adjacent if they have a common side or a common point. In this case, each pixel has eight adjacent pixels. Examples of closed lines and contours of objects formed by boundary pixels are shown in Fig.6. As follows from the above examples, this representation of contour has significant disadvantages [13].

Fig. 7a shows a closed line drawn according to the rules of 4-neighborhoods. As a result of the use of 4-neighborhoods, the contour line divides the image field not into two areas, as it should be, but into three.

A closed simply connected line drawn according to the rules of 8-neighborhoods is shown on Fig. 7b. But, due to 8 neighborhoods, the space inside the closed line and outside of it is not divided: there are connections between pixels inside and outside the line.

Fig. 7c shows an attempt to construct a closed contour of the object using the boundary pixels adjacent to the pixels of the object according to the rules of 4-neighborhoods: the contour line is not closed.

Fig. 7d illustrates the case of constructing a closed contour of an object using boundary pixels adjacent to the pixels of the object according to the 8-neighborhood rules: the contour line is not simply connected.

A curved line in continuous space, as follows from its definition, has no thickness. That is, each of the infinitely large set of points that form a line is an infinitesimal value. This also applies to the closed curve of the object contour line (boundary). At the same time, traditionally used representations of a curved line in a discrete space assume that the curve consists of minimal elements of this space – points. But the point in this case corresponds to a pixel – the minimum element of the image that has finite dimensions. It is this difference that is the reason for the



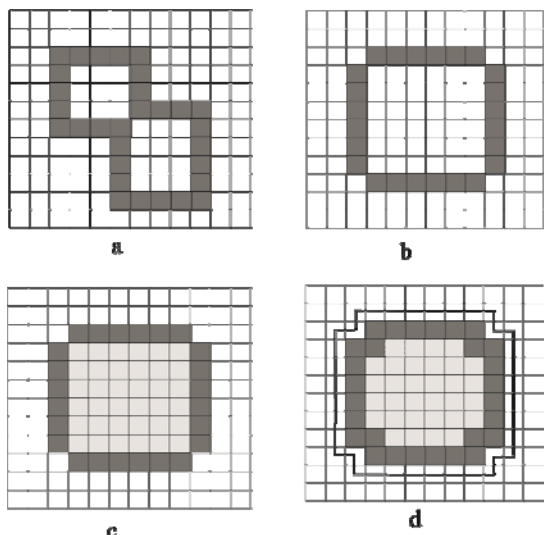


Figure 7 – Closed lines formed by pixels: problems and paradoxes

above paradoxes of representing lines as a sequence of pixels. That is, the correct representation of initially continuous images in discrete space is possible using the theory of cell complexes.

In the primary visual (striate) cortex, neurons were found that generate signals – responses to extended pieces at the border of contrast areas of the visual field, and those and only those neurons are excited whose receptive fields match the border and their orientation match the orientation of the corresponding sections of the contrast area border. That is, the excited neurons respond to the boundary segments of the straight line in certain orientations, which can be considered as I-cells if the contrast area is considered as a cellular complex. That is, the representation of an image as a cellular complex can be considered as an approximation to the implementation of the mechanisms of visual perception. It can also be assumed that the signals of a curved line are formed by pairs of segments of the corresponding directions, as proposed in this paper.

## 6 CONCLUSIONS

The paper considers the arc of the curve as a structural element of the image, more precisely as a structural element of the interest object contour in the image, and the image is presented as a cellular complex.

**The scientific novelty** is that the arc of a digital curve is defined in the discrete space of a digital image, in contrast to the known definitions of continuous curves, which are oriented to use in a continuous space.

**The practical significance** is that the interest objects contours are presented as sequences of line segments and arcs of digital curves. This representation of the object contour does not depend on affine transformations, such as position in the field of view and rotation, which greatly simplifies image processing.

**Prospects for further research** are as follows. The successful result of object contour recognition in the form of a sequence of straight line segments and arcs of a digi-

tal curve depends on the choice of  $d$  – the resolution value when sampling the image. The task of determining the most appropriate resolution for a particular image has not been solved. It is all the more possible that different parts of the same image must be processed at different resolutions. Therefore, the recognition of line segments and arcs of curves using variable resolution will be considered in subsequent publications.

## ACKNOWLEDGEMENTS

The work is supported by the state budget scientific research project of the Institute of Mathematical Machines and Systems Problems “Structural methods of processing cyclic biomedical signals and cloud services based on them” (state registration number 0121U110584).

## REFERENCES

1. Pavlidis T. Algorithms for Graphics and Image Processing. Berlin, Springer-Verlag, 1982, 400 p.
2. Gonzalez R. C., Woods R. E., Eddins S. L. Digital Image Processing using MATLAB. New York, Pearson Education, 2004, 616 p.
3. Pratt W. K. Digital Image Processing. New York, John Wiley & Sons, Inc, 1982, 738 p.
4. Schlesinger M., Hlavac V. Ten Lectures on Statistical and Structural Pattern Recognition. Dordrecht / Boston / London, Computational Imaging and Vision Kluwer Academic Publishers, 2002. 520 p.
5. Ivakhnenko A. G., Lapa V. G. Cybernetics and Forecasting Techniques. New York, American Elsevier Publishing Company, 1967, 168 p.
6. LeCun Y., Boser B., Denker J. S., Henderson D., Howard R. E., Hubbard W., Jackel L. D. Backpropagation Applied to Handwritten Zip Code Recognition, *Neural Computation*, 1989, Vol. 1, No. 4, pp. 541–551. doi:10.1162/neco.1989.1.4.541. S2CID 41312633.
7. Maitra D. S., Bhattacharya U., Parui S. K. CNN based common approach to handwritten character recognition of multiple scripts, *13th International Conference on Document Analysis and Recognition (ICDAR): 23-26 August 2015: proceedings*. Tunis, IEEE 2015, pp. 1021–1025. doi:10.1109/ICDAR.2015.7333916. ISBN 978-1-4799-1805-8. S2CID 25739012.
8. Fu K. S. Syntactic Methods in Pattern Recognition. New York and London, Academic Press, 1974, 511 p.
9. Aleksandrov P. S. Combinatorial Topology. Rochester, Graylock Press, 1956, 656 p.
10. Kovalevsky V. Finite Topology as Applied to Image Analysis, *Computer Vision, Graphics and Image Processing*, 1989, Vol. 46, No. 2, pp. 141–161.
11. Hubel D. H. Eye, brain, and vision. New York, Scientific American Library, Distributed by W.H. Freeman, 1988, 240 p.
12. Berg G. O., Julian W., Mines R., Richman F. The constructive Jordan curve theorem, *Rocky Mountain Journal of Mathematics*, 1975, Vol. 5, № 2, pp. 225–236. DOI: 10.1216/RMJ-1975-5-2-225, ISSN 0035-7596, MR 0410701
13. Dovgoshey O., Martio O., Ryazanov V., Vuorinen M. The Cantor function, *Expositiones Mathematicae. Elsevier BV*, 2006, Vol. 24, № 1, pp. 1–37. DOI: 10.1016/j.exmath.2005.05.002. ISSN 0723-0869. MR 2195181

14. Alexandrov A. D., Reshetnyak Yu. G. General theory of irregular curves, *Mathematics and its Applications (Soviet Series)*, 29. Kluwer, Academic Publishers Group, Dordrecht, 1989, 288 p. ISBN: 90-277-2811-9
15. Schlesinger M. I. *Mathematical Tools of Picture Processing*. Kyiv, Naukova Dumka, 1989, 117 p.
16. Kovalevsky V. A. Applications of Digital Straight Segments to Economical Image Encoding, *7<sup>th</sup> International Workshop, DGCI'97*. Montpellier, France, December 3–5 1997, proceedings, Springer 1997, pp. 51–62.

Accepted 09.11.2022.  
Received 05.01.2023.

УДК 004.93

### ДУГА КРИВОЇ ЯК СТРУКТУРНИЙ ЕЛЕМЕНТ ЗОБРАЖЕННЯ, ЩО МАЄ БУТИ РОЗПІЗНАНЕ

**Калмиков В. Г.** – канд. техн. наук, старший науковий співробітник Інституту проблем математичних машин і систем, Київ, Україна.

**Шарипанов А. В.** – канд. техн. наук, старший науковий співробітник Інституту проблем математичних машин і систем, Київ, Україна.

**Вишневецький В. В.** – канд. техн. наук, провідний науковий співробітник Інституту проблем математичних машин і систем, Київ, Україна.

#### АНОТАЦІЯ

**Актуальність.** Пропонована стаття стосується галузі обробки візуальної інформації в комп'ютерному середовищі, а саме визначення параметрів об'єкта інтересу на зображенні, зокрема контуру об'єкта інтересу. У більшості випадків контур об'єкта інтересу є однозв'язна послідовність дуг кривих.

**Мета.** Мета і предмет дослідження – знайти і запропонувати таке визначення дуги цифрової кривої, як найважливішого елемента контуру об'єкта в розпізнаваному образі, яке не суперечить сучасним нейрофізіологічним уявленням про зорове сприйняття, і розпізнати контур об'єкта як послідовність дуг цифрових кривих.

В якості **методу** використовується подання зображення у вигляді структурної моделі, одним із структурних елементів якої є контур об'єкта, що складається з цифрових дуг кривих. Також зображення розглядається як клітинний комплекс, що відповідає сучасним уявленням про зорове сприйняття людини.

**Результати.** Запропоновано нове визначення дуги цифрової кривої як послідовності відрізків цифрових прямих, що не суперечить сучасним уявленням нейрофізіології. На відміну від відомих визначень дуги кривої, запропоноване визначення дуги цифрової кривої дає можливість визначити початкову та кінцеву точки дуги. За описом контуру об'єкта як однозв'язної замкнутої послідовності відрізків пропонується побудувати опис контуру як послідовності дуг цифрових кривих.

**Висновки.** Використання запропонованого визначення дуги цифрової кривої при обробці зображень дає змогу розпізнати контур об'єкта на зображенні та представити його у формі, наближеній до зорового сприйняття. Для досягнення найкращих результатів рекомендується використовувати змінну роздільну здатність в алгоритмах обробки зображень.

**КЛЮЧОВІ СЛОВА:** зображення, контур, дуга кривої, відрізок прямої, клітинний комплекс, нейрони, рецептивне поле.

#### ЛІТЕРАТУРА

1. Pavlidis T. *Algorithms for Graphics and Image Processing* / T. Pavlidis. – Berlin : Springer-Verlag, 1982. – 400 p.
2. Gonzalez R. C. *Digital Image Processing using MATLAB* / R. C. Gonzalez, R. E. Woods, S. L. Eddins. – New York : Pearson Education, 2004. – 616 p.
3. Pratt W. K. *Digital Image Processing* / W. K. Pratt. – New York : John Wiley & Sons, Inc, 1982. – 738 p.
4. Schlesinger M. *Ten Lectures on Statistical and Structural Pattern Recognition* / M. Schlesinger, V. Hlavac. – Dordrecht / Boston / London: Computational Imaging and Vision Kluwer Academic Publishers, 2002. – 520 p.
5. Ivakhnenko A.G. *Cybernetics and Forecasting Techniques* / A.G. Ivakhnenko, V. G. Lapa. – New York : American Elsevier Publishing Company, 1967. – 168 p.
6. Backpropagation Applied to Handwritten Zip Code Recognition / [Y. LeCun, B. Boser, J.S. Denker et al.] // *Neural Computation*. – 1989. – Vol. 1, № 4. – P. 541–551. DOI: 10.1162/neco.1989.1.4.541. S2CID 41312633.
7. Maitra D. S. CNN based common approach to handwritten character recognition of multiple scripts / D. S. Maitra, U. Bhattacharya, S. K. Parui // *13th International Conference on Document Analysis and Recognition (ICDAR): 23–26 August 2015: proceedings*. – Tunis: IEEE 2015. – P. 1021–1025. DOI: 10.1109/ICDAR.2015.7333916. ISBN 978-1-4799-1805-8. S2CID 25739012.
8. Fu K. S. *Syntactic Methods in Pattern Recognition* / K. S. Fu. – New York and London: Academic Press, 1974. – 511 p.
9. Aleksandrov P. S. *Combinatorial Topology* / P. S. Aleksandrov. – Rochester : Graylock Press, 1956. – 656 p.
10. Kovalevsky V. *Finite Topology as Applied to Image Analysis* / V. Kovalevsky // *Computer Vision, Graphics and Image Processing*. – 1989. – Vol. 46, No. 2. – P. 141–161.
11. Hubel D. H. *Eye, brain, and vision* / D. H. Hubel. – New York : Scientific American Library, Distributed by W.H. Freeman, 1988. – 240 p.
12. The constructive Jordan curve theorem / [G. O. Berg, W. Julian, R. Mines, F. Richman] // *Rocky Mountain Journal of Mathematics*. – 1975. – Vol. 5, № 2. – P. 225–236. DOI: 10.1216/RMJ-1975-5-2-225, ISSN 0035-7596, MR 0410701
13. The Cantor function / [O. Dovgoshey, O. Martio, V. Ryzanov, M. Vuorinen] // *Expositiones Mathematicae*. Elsevier BV. – 2006. – Vol. 24, № 1. – P. 1–37. DOI:10.1016/j.exmath.2005.05.002. ISSN 0723-0869. MR 2195181
14. Alexandrov A. D. General theory of irregular curves / A. D. Alexandrov, Yu. G. Reshetnyak // *Mathematics and its Applications (Soviet Series)*, 29. – Kluwer Academic Publishers Group, Dordrecht. – 1989. – 288 p. ISBN: 90-277-2811-9
15. Schlesinger M. I. *Mathematical Tools of Picture Processing* / M. I. Schlesinger. – Kyiv: Naukova Dumka, 1989. – 117 p.
16. Kovalevsky V. A. Applications of Digital Straight Segments to Economical Image Encoding / V. A. Kovalevsky // *7<sup>th</sup> International Workshop, DGCI'97 – Montpellier, France, December 3–5 1997 – proceedings*. – Springer 1997. – P. 51–62.

## APPLICATION OF TWO-DIMENSIONAL PADÉ-TYPE APPROXIMATIONS FOR IMAGE PROCESSING

**Olevskiy V. I.** – Dr. Sc., Professor, Professor of the Information Technology and Computer Engineering Department, Dnipro University of Technology, Dnipro, Ukraine.

**Hnatushenko V. V.** – Dr. Sc., Professor, Head of the Department of Information Technology and Computer Engineering Department, Dnipro University of Technology, Dnipro, Ukraine.

**Korotenko G. M.** – Dr. Sc., Associate Professor, Professor of the Information Technology and Computer Engineering Department, Dnipro University of Technology, Dnipro, Ukraine.

**Olevska Yu. B.** – PhD, Associate Professor, Associate Professor of the Applied Mathematics Department, Dnipro University of Technology, Dnipro, Ukraine.

**Obydennyi Ye. O.** – Assistant Professor of the of Information Technology and Computer Engineering Department, Dnipro University of Technology, Dnipro, Ukraine.

### ABSTRACT

**Context.** The Gibbs phenomenon introduces significant distortions for most popular 2D graphics standards because they use a finite sum of harmonics when image processing by expansion of the signal into a two-dimensional Fourier series is used in order to reduce the size of the graphical file. Thus, the reduction of this phenomenon is a very important problem.

**Objective.** The aim of the current work is the application of two-dimensional Padé-type approximations with the aim of elimination of the Gibbs phenomenon in image processing and reduction of the size of the resulting image file.

**Method.** We use the two-dimensional Padé-type approximants method which we have developed earlier to reduce the Gibbs phenomenon for the harmonic two-dimensional Fourier series. A definition of a Padé-type functional is proposed. For this purpose, we use the generalized two-dimensional Padé approximation proposed by Chisholm when the range of the frequency values on the integer grid is selected according to the Vavilov method. The proposed scheme makes it possible to determine a set of series coefficients necessary and sufficient for construction of a Padé-type approximation with a given structure of the numerator and denominator. We consider some examples of Padé approximants application to simple discontinuous template functions for both formulaic and discrete representation.

**Results.** The study gives us an opportunity to make some conclusions about practical usage of the Padé-type approximation and about its advantages. They demonstrate effective elimination of distortions inherent to Gibbs phenomena for the Padé-type approximant. It is well seen that Padé-type approximant is significantly more visually appropriate than Fourier one. Application of the Padé-type approximation also leads to sufficient decrease of approximants' parameter number without the loss of precision.

**Conclusions.** The applicability of the technique and the possibility of its application to improve the accuracy of calculations are demonstrated. The study gives us an opportunity to make conclusions about the advantages of the Padé-type approximation practical usage.

**KEYWORDS:** Padé-type approximants, Gibbs phenomenon, size of the image file.

### ABBREVIATIONS

DCT is a discrete cosine transform;

2D DCT is a two-dimensional discrete cosine transform;

### NOMENCLATURE

$L_2 [ ]$  is a metric Hilbert space with a square measure;

$a_i$  are minimal levels of variables;

$b_i$  are maximum levels of variables;

$x_i$  are complex variables on the interval  $(a_i, b_i)$ ;

$f()$  is an arbitrary function in the space under consideration;

$a_{kp}$  are coefficients of harmonics for the 2D Fourier series of  $f$ ;

$B_i$  are countable sets of basic functions;

$e_{ik}$  are basic functions;

$B$  is a basis of the space;

$P [ ]$  is a two-dimensional Padé approximant;

$m_i$  are maximum powers of  $x_1$  for numerator and denominator of  $P$  ;

$n_i$  are maximum powers of  $x_2$  for numerator and denominator of  $P$  ;

$S$  is a two-dimensional power series;

$GS$  is a generalized power series;

$GP_{GS}$  is a Padé-type functional;

$N$  is a number of harmonics for the Fourier series in  $x_1$  direction, and the maximum power of  $x_i$  in both numerator and denominator of  $P$  in the Chisholm approximation;

$M$  is a number of harmonics for the Fourier series in  $x_2$  direction;

$\lambda_i$  are frequencies of 2D Fourier series;

$n_F$  is a number of parameters in the Fourier series;

$n_P$  is a number of parameters in the Padé-type approximation;

$F$  is a 2D DCT for  $f$ ;

$f_{N,N}$  is the Chisholm approximation of  $N$ -th order;

$p$  is a matrix of the Chisholm approximation numerator coefficients;

$p_{kp}$  are elements of  $p$ ;

$q$  is a matrix of the Chisholm approximation denominator coefficients;

$q_{kp}$  are elements of  $q$ .

## INTRODUCTION

Image processing by expansion of the signal into a two-dimensional Fourier series in order to reduce the size of the graphical file often leads to significant image distortion, in particular, due to the Gibbs effect. The Gibbs phenomenon is the property of the one-dimensional Fourier series which manifests in the decomposition of a discontinuous periodic function when they are truncated and a finite number of members are used [1–3]. In the case of truncation a distortion occurs near the discontinuity points which cannot be eliminated by increasing the finite number of terms of the series. In two-dimensional case the Gibbs phenomenon significantly reduces the quality of the processed images for most popular graphic standards because they use a finite sum of harmonics. Distortion occurs on the borders of sharp contrast change and leads to the appearance of false optical shadows. It negatively influences the analysis quality when processing the results of x-ray and sonar studies.

**The object of study** is the image processing.

**The subject of study** is the generalized sum of a two-dimensional Fourier series obtained by image processing using the fractional rational approximants.

**The purpose of the work** is the application of the two-dimensional Padé-type approximations for the purpose of elimination of the Gibbs phenomenon for image processing and the reduction of the size of the resulting image file.

## 1 PROBLEM STATEMENT

Let the original monochrome image be given in the form of a two-dimensional function of tone  $f$  or a set of its values for rectangle  $(a_2, b_2) \times (a_2, b_2)$ . The aim is to find an appropriate two-dimensional Padé-type approximation  $P[m_1, n_1 / m_2, n_2](x_1, x_2)$  to eliminate the Gibbs phenomenon for image processing and reduce the size of an image file.

## 2 REVIEW OF THE LITERATURE

The Gibbs phenomenon also exists in the two-dimensional case, and it significantly reduces the quality of images processing for most popular graphic standards because they use a finite sum of harmonics [2]. Distortion occurs on the borders of sharp contrast change and leads to the appearance of false optical shadows [3, 4]. It is detrimental for the analysis quality when processing the results of the x-ray and sonar studies. Gibbs phenomenon is a type of MRI artifact, which leads to a series of lines in the MRI image parallel to abrupt and intense changes in the object, such as the CSF-spinal cord and the skull-brain interface [4, 5].

The theory of approximation of mathematical physics functions is the most rapidly developing field of mathematics [1, 6, 7]. Traditionally, only the approximation techniques which use polynomials [1, 8, 9] and trigonometric functions [6, 7, 10] are considered. The most successful type of such approximation techniques is the approximation by fractional rational functions [11, 12]. The interest in the theory of fractional-rational approximations has been steadily increasing due to their wide application in various studies in the field of theoretical physics, applied mechanics, geophysics, etc. [8, 11, 12] due to them allowing for a generalized summation of series and the extension of the function to be approximated into the meromorphic domain.

Recently, great attention has been given to the expansion of the classical theory of approximation by fractional-rational functions to various types of basis functions and different methods of constructing approximants – to Padé-type approximations [12–19]. A special choice of the constructing method for the approximation makes it possible in many cases to achieve a significant improvement in the useful properties of the approximants for certain distinct classes of functions [12–14].

We suggest the application of the two-dimensional Padé-type approximants method which we have developed earlier [12] for the purpose of reduction of the Gibbs phenomenon in the harmonic two-dimensional Fourier series.

## 3 MATERIALS AND METHODS

Let's provide some basic principles of construction of the Padé-type approximants for the harmonic two-dimensional Fourier series as a subset of power series. According to our approach [12], we consider the separable space  $L_2[(a_2, b_2) \times (a_2, b_2)]$  of two-dimensional complex functions  $f(x_1, x_2)$ , which are integrable on this rectangle. The boundaries of the rectangle can be finite or infinite. We can choose a countable set of functions  $B_1 = \{e_{1k}, k = \overline{1, \infty}\}$  and  $B_2 = \{e_{2j}, j = \overline{1, \infty}\}$  as a basis of space with respect to individual coordinates of the form

$$B = \{e_{1k} e_{2j}, k = \overline{1, \infty}, j = \overline{1, \infty}\}. \quad (1)$$

The basic functions in the case of trigonometric functions can be represented in the form

$$e_{nk} = \left( e^{ix_n} \right)^k = e^{ikx_n}, \quad n = 1, 2. \quad (2)$$

The expansion of an arbitrary function in the space under consideration with respect to the basis (2) can be regarded as a two-dimensional generalized power series of the form

$$f = \sum_{k, p=1}^{\infty} a_{kp} (e_{11})^k (e_{21})^p. \quad (3)$$

In [12] we have proposed the definition of the functional of the Padé-type in following form.

Definition. Suppose a two-dimensional power series  $S = \sum_{k,p=1}^{\infty} a_{kp} (x_1)^k (x_2)^p$  of complex variables  $x_1$  and  $x_2$  and the associated Padé approximant  $P[m_1, n_1 / m_2, n_2](x_1, x_2)$  in the proper sense are given. The Padé-type functional  $GP_{GS}[m_1, n_1 / m_2, n_2](f_1, f_2)$  associated with the given generalized power series  $GS = \sum_{k,p=1}^{\infty} a_{kp} (f_1)^k (f_2)^p$  for the complex functions of these variables is defined as

$$GP_{GS}[m_1, n_1 / m_2, n_2](f_1, f_2) = P[m_1, n_1 / m_2, n_2](x_1, x_2) \Big|_{x_1=f_1, x_2=f_2}. \quad (4)$$

The following sequence details the construction process:

1. The types of bases for the individual variables  $B_1, B_2$  and the basis of the space  $B(1)$  are chosen.
2. The function  $f$  to be approximated is represented in the form (3).
3. For the power series of two complex variables  $x_1$  and  $x_2$  with coefficients coinciding with (3), a Padé approximant  $P[m_1, n_1 / m_2, n_2](x_1, x_2)$  is constructed in the proper sense.
4. A substitution of basis functions into a functional of Padé-type (4) is performed.

The proposed scheme makes it possible to determine the set of coefficients of a series that is necessary and sufficient for the construction of the Padé-type approximant with a given structure of the numerator and denominator.

If we have a function  $f(x_1, x_2)$  which represents brightness of monochrome image point in the range of  $[0,1]$  on the rectangle  $(a_2, b_2) \times (a_2, b_2)$  in form of truncated Fourier series

$$f(x_1, x_2) \approx \sum_{m=0}^M \sum_{n=0}^N f_{mn} \cos(m\lambda_1 x_1) \cos(n\lambda_2 x_2),$$

and it's Padé approximant  $P[m_1, n_1 / m_2, n_2](x_1, x_2)$ , then to obtain cosine part of two-dimensional exponent, we use the following equality:

$$\cos x \cos y = \frac{1}{2} \operatorname{Re} \left( e^{ix} e^{iy} + e^{ix} e^{-iy} \right).$$

As was considered in [18], in this case

$$f(x_1, x_2) \approx \frac{1}{2} \operatorname{Re} \left[ P(x_1, x_2) + P(x_1, -x_2) \right]. \quad (5)$$

If the image is stored as a two-dimensional array of points (for example, as a bmp file), discrete Fourier transform procedures can be used to implement Padé approximation [2]. DFT is the basis of many image and video compression algorithms, especially the basic jpeg and mpeg standards for compressing both still and video images.

The input image is decomposed into spectral components using a two-dimensional discrete cosine transform (2D DCT). 2D DCT can be calculated by applying a one-dimensional DCT algorithm for each row or column of a two-dimensional matrix of the input signal, since DCT is a separable function. The direct two-dimensional DCT of a  $M \times N$  matrix of a two-dimensional signal  $f(x,y)$  can be written as

$$F(u, v) = \frac{2}{\sqrt{MN}} C(u) C(v) \times \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} f(n, m) \cos\left(\frac{\pi(2n+1)u}{2N}\right) \cos\left(\frac{\pi(2m+1)v}{2M}\right). \quad (6)$$

where

$$C(u) = \begin{cases} \frac{1}{\sqrt{2}}, & \text{for } u = 0, \\ 1, & \text{for } u \neq 0. \end{cases}$$

Coefficients of a truncated cosine Fourier series for  $f(x_1, x_2)$  can be obtained as the values of  $F(u, v)$  (6), divided by the step.

Next, the range of the frequency values on the integer grid is selected according to the Vavilov method [13]. The size of this range directly determines the number of equations that must be generated.

For this purpose, we use the generalized two-dimensional Padé approximation for case  $N=M$  proposed by Chisholm [11]. If  $f(x,y)$  is a function of two variables with a two-dimensional expansion into a power series of the form

$$f(x, y) = \sum_{k,p=1}^{\infty} a_{kp} x^k y^p.$$

then the  $N$ -th Chisholm approximation can be written as

$$f_{N,N}(x, y) = \frac{\sum_{k,p=1}^N p_{kp} x^k y^p}{\sum_{k,p=1}^N q_{kp} x^k y^p}.$$

If we use a set of power values bounded by a right triangle with the axes being its legs, then the coefficients  $p_{kp}$  and  $q_{kp}$  can be calculated with the help of the following equations

$$\sum_{\sigma=0}^{\gamma} \sum_{r=0}^{\delta} q_{\sigma r} a_{\gamma-\sigma, \delta-r} = P_{\gamma \delta},$$

$$(\gamma, \delta = 0, 1, \dots, 2N, 1 \leq \delta + \gamma \leq 2N),$$

$$\sum_{\sigma=0}^{\gamma} \sum_{r=0}^{\delta} (q_{\sigma r} a_{\gamma-\sigma, \delta-r} + q_{r \sigma} a_{\delta-r, \gamma-\sigma}) = 0,$$

$$(\gamma = 1, 2, \dots, 2N, \delta + \gamma = 2N),$$

$$p_{00} = 1.$$

The solution of this system of equations are matrixes of coefficients p and q.

The algorithm for compression of two-dimensional signals using two-dimensional discrete cosine transformation is shown in Fig. 1.

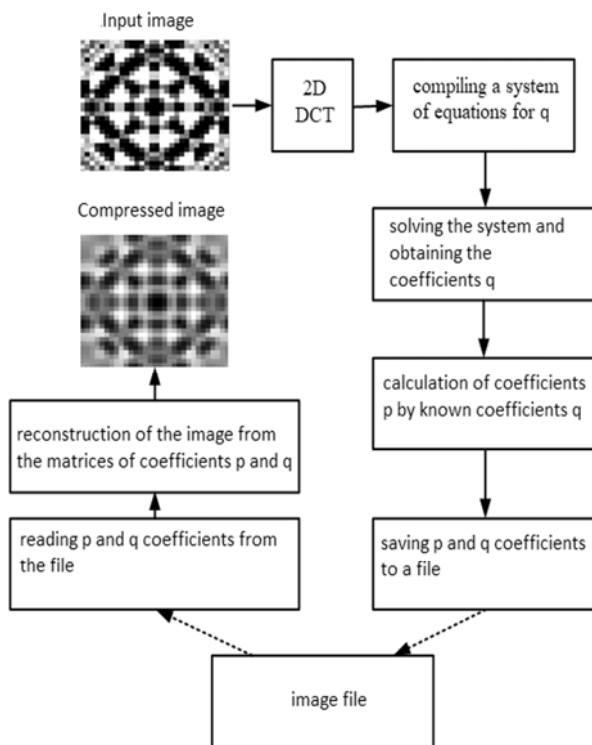


Figure 1 – Example of training samples formation from the original sample for the Fischer problem in the space

#### 4 EXPERIMENTS

Here we consider some examples of Padé approximants application to simple discontinuous template functions for both formulaic and discrete representation. In general case images could be also represented using such series [2].

Let's consider a periodic template function with the period of  $2\pi$  in the form

$$f(x_1, x_2) = \begin{cases} 1, & x_1^2 + x_2^2 \leq \pi, \\ 0, & x_1^2 + x_2^2 > \pi. \end{cases} \quad (7)$$

This function is symmetrical about both  $x_1$  and  $x_2$  axes. Thus, the most suitable way of its approximation is a truncated Fourier series in the form

$$f(x_1, x_2) \approx \sum_{m=0}^4 \sum_{n=0}^4 f_{mn} \cos(mx_1) \cos(nx_2). \quad (8)$$

We have applied Padé-type approximation to (6) with structure  $[2, 2/2, 2]$ , using subset of previously estimated Fourier coefficients (6) and no additional information. To obtain the cosine part of the two-dimensional exponent, we use the converse transformation (5) for Padé-type approximant  $P[2, 2/2, 2](x_1, x_2)$  to obtain the desired real approximation. Transformation which is used for cosine series is the same as the one used in radio physics [10].

We also consider the six monochrome symmetrical bitmap test images from the digital library [20]. These were also used as input data and can be seen in Table 1. In the case of an asymmetric input signal, the image can be artificially expanded to a symmetrical one.

Table 1 – Initial and restored images with the number of harmonics  $N=8$

No	Initial image	Image Padé-type approximant	Image with compression
1			
2			
3			
4			
5			
6			

When using the jpeg standard, insignificant decomposition coefficients are excluded in order to reduce the file size. This procedure was used by processing the results of 2D DCP of the input image, considering the rapid decline of harmonic amplitudes, and it was this image that was used for comparison with the quality of the one compressed by the proposed method. Standardized root mean square error and normalized mean absolute error were used as comparison criteria.

### 5 RESULTS

Two-dimensional grayscale images used the template periodic function (7). The resulting truncated Fourier series (8) and its Padé-type approximant are represented on Fig. 2a, 2b and 2c respectively. For the Fourier series the image demonstrates distortions inherent to Gibbs phenomena, and their effective absence for the Padé-type approximant. It is well seen that the Padé-type approximant is much more visually appropriate than Fourier one.

In order to assess the accuracy of the Fourier series method of and the Padé-type approximation, Fig. 3 presents their one-dimensional sections for comparison with the template function. This also demonstrates the advantage of the Padé approximation.

The size of the area on the integer grid was chosen in the range between 2 and 8, while the number of coefficients by which the reconstruction of the compressed image was performed, was gradually increased. For the compressed image this number (and, therefore, the volume of the graphic file) was approximately half of the value. Two extreme cases are schematically shown in Fig. 4.

A subjective assessment of the quality of the restored image can be obtained from Tab. 1, which shows the input and the reconstructed images.

### 6 DISCUSSIONS

Analyzing the graphs of the mean square error, one can notice a sharp decrease in the mean square error when the image approaches the psycho-visual similarity to the

original. The results of the criteria calculations showed that each type of image has its own lower limit when the reconstructed image visually correlates with the initial one (Table 2).

Table 2 – The number of harmonics for minimum error

image No	1	2	3	4	5	6
the number of harmonics $N$	7	6	4	4	5	8

The study makes it possible to draw some conclusions about the practical use of the Padé-type approximation method and its advantages.

First of all, one can see the low level of noise for the Padé approximation (Fig. 5b) compared to the Fourier series for the cosine (Fig. 5a).

Secondly, the use of Padé-type approximation leads to a sharp decrease in the number of approximant parameters without the loss of accuracy (and even with its increase). Indeed, when using Fourier series with an equal number of  $N$  harmonics in both directions, the following number of parameters  $n_F = N^2$  is obtained. Using the Padé-type approximation with the same powers of the numerator and denominator equal to  $N/2$  gives the following number of parameters:

$$n_P = 2 \left( \frac{N}{2} \right)^2 - 1 = \frac{N^2}{2} - 1.$$

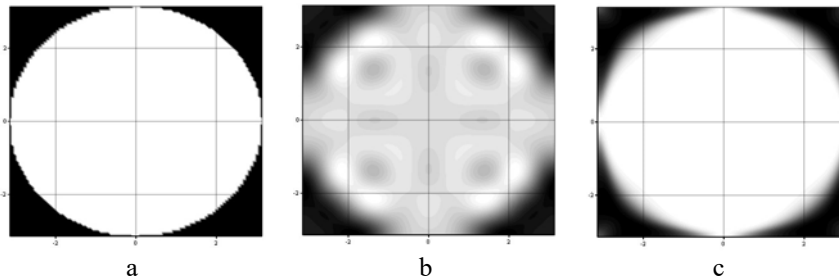


Figure 2 – Images of a – the template function, b – Fourier series  $4 \times 4$ , c – Padé Approximant  $[2, 2 / 2, 2]$ .

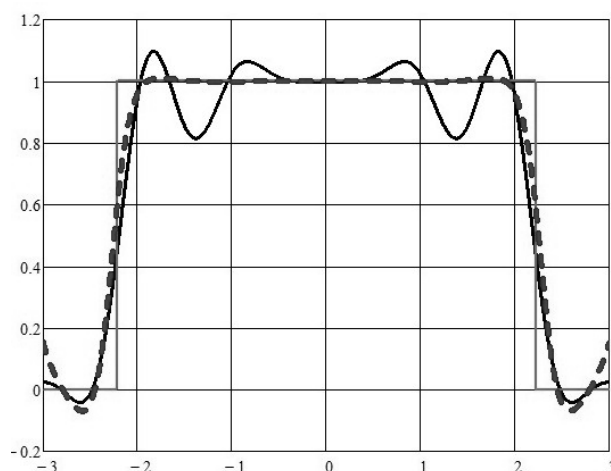


Figure 3 – Cross-sections of approximants and template along line  $x_1 = x_2$ . Grey line – the template, black – Fourier series, dashed line – Padé approximant.

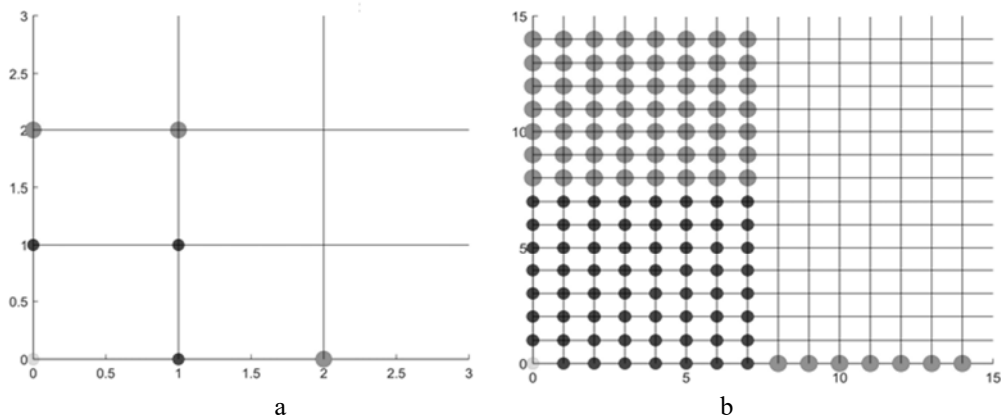


Figure 4 – An integer grid of power values for the two extreme cases: a –  $N=2$ , b –  $N=8$ .

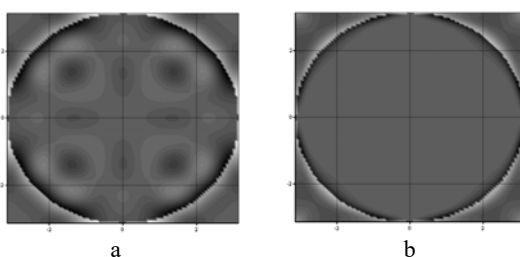


Figure 5 – Distortions between template functions and their approximations for a – Fourier cosine series, b – Padé approximation of cosine series

Thus, the number of the parameters is more than halved:

$$\frac{n_F}{n_P} > 2.$$

This is very important for the purpose of saving the images in digital signal processing and can provide a theoretical basis for building a new effective image format similar to the well-known jpeg format [1, 2, 10].

### CONCLUSIONS

An important problem of applied mathematics is solved in order to reduce the Gibbs phenomenon for the harmonic two-dimensional Fourier series.

**The scientific novelty** of obtained results is that they demonstrate effective absence of distortions inherent to Gibbs phenomena for the Padé-type approximant. It is well seen that the Padé-type approximant is much more visually appropriate than Fourier one. Application of the Padé-type approximation also leads to the sufficient decrease of the approximants' parameter number without the loss of precision.

**The practical significance** of obtained results is that the software implementing the proposed method is fit for practical use along with the estimation of the appropriate application conditions.

**Prospects for further research** are to study the proposed method as a theoretical basis for building a new

effective image format similar to the well-known jpeg format.

### ACKNOWLEDGEMENTS

The work is supported by the state budget scientific research project of Dnipro University of Technology “Models and information technologies of data processing and analysis in complex computer systems and networks” (state registration number 0121U114523).

### REFERENCES

1. Timan A. F. Theory of approximation of functions of a real variable. New York, MacMillan, 1963, 631 p. <https://doi.org/10.1016/c2013-0-05307-8>
2. Mitra S. K. Digital Signal Processing: A Computer-Based Approach. New York, McGraw-Hill, 2001, 866 p. [https://doi.org/10.1016/s0026-2692\(98\)00072-x](https://doi.org/10.1016/s0026-2692(98)00072-x)
3. Helmbert G. Localization of a Corner-Point Gibbs Phenomenon for Fourier Series in Two Dimensions, *Journal of Fourier Analysis and Applications*, 2002, Vol. 8(1), pp. 29–42. DOI: 10.1007/s00041-002-0002-9
4. Archibald R. and Gelb A. A method to reduce the Gibbs ringing artifact in MRI scans while keeping tissue boundary integrity, *IEE Transactions of Medical Imaging*, 2002, Vol. 21(4), pp. 305–319. DOI: 10.1109/TMI.2002.1000255
5. Veraart J., Fieremans E., Jelescu I. O., Knoll F., and Novikov D. S. Gibbs ringing in diffusion MRI, *Magn. Reson. Med.*, 2016, Vol. 76, pp. 301–314. DOI: 10.1002/mrm.25866
6. Serov V. Fourier Series, Fourier Transform and Their Applications to Mathematical Physics. New York, Springer International Publishing, 2017, 534 p. DOI: 10.1007/978-3-319-65262-7.



7. Maggioli F., Melzi S., Ovsjanikov M., Bronstein M. M., Rodolà E. Orthogonalized Fourier polynomials for signal approximation and transfer, *Computer Graphics Forum*, 2021, Vol. 40(2), pp. 435–447. <https://doi.org/10.1111/cgf.142645>
8. Andrianov I., Awrejcewicz J., Danishevskyy V., Ivankov A. Asymptotic Methods in the Theory of Plates with Mixed Boundary Conditions. New York, John Wiley & Sons, 2014, 288 p. DOI: 201410.1002/9781118725184.
9. Olevska Yu. B., Olevskiy V. I., Olevskiy O. V. Using of fuzzy mathematical models in automated systems for recognition of high molecular substances, *Application of Mathematics in Technical and Natural Sciences: 10th International Conference for Promoting the Application of Mathematics in Technical and Natural Sciences – AMiTaNS'18, Albena, 20–25 June: proceedings*. New York, American Institute of Physics, Melville, NY, 2018, pp. 060003-1–060003-9. (AIP Conference Proceedings, Vol. 2025(1)). <https://doi.org/10.1063/1.5064911>
10. Prots'ko I. O., Kuzminskij R. D., Teslyuk V. M. Efficient computation of the integer DCT-II for compressing images, *Radio Electronics, Computer Science, Control*, 2019, No. 2, pp. 151–157. <https://doi.org/10.15588/1607-3274-2019-2-16>
11. Baker J. A., Jr. and Graves-Morris P. Padé approximants. New York, Cambridge University Press, 1996, 746 p. <https://doi.org/10.1017/cbo9780511530074>
12. Andrianov I. V., Olevskiy V. I., Shapka I. V., Naumenko T. S. Technique of Padé-type multidimensional approximations application for solving some problems in mathematical physics, *Application of Mathematics in Technical and Natural Sciences: 10th International Conference for Promoting the Application of Mathematics in Technical and Natural Sciences – AMiTaNS'18, Albena, 20–25 June, 2018: proceedings*. New York, American Institute of Physics, Melville, NY, 2018, pp. 040002-1–040002-9. (AIP Conference Proceedings, Vol. 2025 (1)). DOI: 10.1063/1.5064886
13. Bosuwan N., López Lagomasino G. Inverse Theorem on Row Sequences of Linear Padé-orthogonal Approximation, *Comput. Methods Funct. Theory*, 2015, Vol. 15, pp. 529–554. <https://doi.org/10.1007/s40315-015-0121-3>
14. Labych Yu. A., Starovoitov A. P. Trigonometric Padé approximants for functions with regularly decreasing Fourier coefficients, *Sb. Math*, 2009, Vol. 200(7), pp. 1051–1074. DOI: 10.1070/SM2009v200n07ABEH004027
15. Buslaev V. I., Suetin S. P. On the existence of compacta of minimal capacity in the theory of rational approximation of multi-valued analytic functions, *J. Approx. Theory*, 2016, Vol. 206, pp. 48–67. DOI: 10.1016/j.jat.2015.08.002
16. Sablonniere P. Padé-Type Approximants for Multivariate Series of Functions, *Lecture Notes in Mathematics*, 1984, Vol. 1071, pp. 238–251. <https://doi.org/10.1007/bfb0099622>
17. Kida S. Padé-type and Padé approximants in several variables, *Appl. Numer. Math*, 1989/90, Vol. 6, pp. 371–391. [https://doi.org/10.1016/0168-9274\(90\)90027-D](https://doi.org/10.1016/0168-9274(90)90027-D)
18. Olevska Yu. B., Olevskiy V. I., Shapka I. V., and Naumenko T. S. Application of two-dimensional Padé-type approximants for reducing the Gibbs phenomenon, *Application of Mathematics in Technical and Natural Sciences: 11th International Conference for Promoting the Application of Mathematics in Technical and Natural Sciences – AMiTaNS'19, Albena, 20–25 June, 2019: proceedings*. New York, American Institute of Physics, Melville, NY, 2018, pp. 060014-1–060014-8. (AIP Conference Proceedings, Vol. 2164). <https://doi.org/10.1063/1.5130816>
19. Daras N. J. The convergence of Padé-type approximants to holomorphic functions of several complex variables, *Appl. Numer. Math*, 1989/90, Vol. 6, pp. 341–360. [https://doi.org/10.1016/0168-9274\(90\)90025-B](https://doi.org/10.1016/0168-9274(90)90025-B)
20. TESTIMAGES free collection of digital images for testing [Electronic resource]. Access mode: <https://testimages.org/Received00.00.2023>.

Received 04.01.2023.  
Accepted 01.02.2023.

УДК 004.93

## ЗАСТОСУВАННЯ ДВОВИМІРНИХ АПРОКСИМАЦІЙ ТИПУ ПАДЕ ДЛЯ ОБРОБКИ ЗОБРАЖЕНЬ

**Олевський В. І.** – д-р. техн. наук, професор, професор кафедри інформаційних технологій та комп'ютерної інженерії НТУ «Дніпровська політехніка», Дніпро, Україна.

**Гнатушенко В. В.** – д-р. техн. наук, професор, завідувач кафедри інформаційних технологій та комп'ютерної інженерії НТУ «Дніпровська політехніка», Дніпро, Україна.

**Коротенко Г. М.** – д-р. техн. наук, доцент, професор кафедри інформаційних технологій та комп'ютерної інженерії НТУ «Дніпровська політехніка», Дніпро, Україна.

**Олевська Ю. Б.** – канд. фіз.-мат. наук, доцент, доцент кафедри прикладної математики НТУ «Дніпровська політехніка», Дніпро, Україна.

**Обиденний Є. О.** – асистент кафедри інформаційних технологій та комп'ютерної інженерії НТУ «Дніпровська політехніка», Дніпро, Україна.

### АНОТАЦІЯ

**Актуальність.** У двовимірному випадку феномен Гіббса значно погіршує обробку зображень для більшості популярних графічних стандартів, оскільки вони використовують кінцеву суму гармонік коли використовується обробка зображення шляхом розкладання сигналу в двовимірний ряд Фур'є з метою зменшення розміру графічного файлу. Тому зменшення цього явища є дуже важливою проблемою.

**Мета роботи.** Метою роботи є використання двовимірних апроксимацій типу Паде для усунення феномену Гіббса під час обробки зображень та зменшення розміру файлу зображення.

**Метод.** Ми використовуємо метод двовимірних апроксимацій типу Паде, який ми розробили раніше, щоб зменшити феномен Гіббса для гармонійного двовимірного ряду Фур'є. Запропоновано визначення функціонала типу Паде. Для цього використовується узагальнена двовимірна апроксимація Паде, запропонована Чізхольмом, при цьому діапазон значень частоти на цілочисельній сітці вибирається за методом Вавілова. Запропонована схема дає змогу визначити набір коефіцієнтів

ряду, необхідний і достатній для побудови апроксимації типу Паде із заданою структурою чисельника та знаменника. Розглядаються деякі приклади застосування апроксимацій Паде до простих розривних шаблонних функцій як для аналітичного, так і для дискретного представлення.

**Результати.** Наше дослідження дає можливість зробити деякі висновки щодо практичного використання апроксимації типу Паде та її переваг. Вони демонструють практичну відсутність спотворень для апроксиманти типу Паде, властивої саме явищам Гіббса. Добре видно, що апроксимація типу Паде є набагато зручнішою візуально, ніж апроксимація Фур'є. Використання апроксимації типу Паде також призводить до значного зменшення кількості параметрів апроксимантів без втрати точності.

**Висновки.** Продемонстровано працездатність методики та можливість її застосування для підвищення точності розрахунків. Дослідження дає можливість зробити висновки про переваги практичного використання апроксимації типу Паде.

**КЛЮЧОВІ СЛОВА:** апроксимації типу Паде, феномен Гіббса, розмір файлу зображення.

#### ЛІТЕРАТУРА

1. Timan A. F. Theory of approximation of functions of a real variable / A. F. Timan. – New York : MacMillan, 1963. – 631 p. <https://doi.org/10.1016/c2013-0-05307-8>
2. Mitra S. K. Digital Signal Processing: A Computer-Based Approach / S. K. Mitra. – New York : McGraw-Hill, 2001. – 866 p. [https://doi.org/10.1016/s0026-2692\(98\)00072-x](https://doi.org/10.1016/s0026-2692(98)00072-x)
3. Helmborg G. Localization of a Corner-Point Gibbs Phenomenon for Fourier Series in Two Dimensions / G. Helmborg // Journal of Fourier Analysis and Applications. – 2002. – Vol. 8(1). – P. 29–42. DOI: 10.1007/s00041-002-0002-9
4. Archibald R. A method to reduce the Gibbs ringing artifact in MRI scans while keeping tissue boundary integrity / R. Archibald and A. Gelb // IEE Transactions of Medical Imaging. – 2002. – Vol. 21(4). – P. 305–319. DOI: 10.1109/TMI.2002.1000255
5. Gibbs ringing in diffusion MRI / [J. Veraart, E. Fieremans, I. O. Jelescu et al.] // Magn. Reson. Med. – 2016. – Vol. 76. – P. 301–314. DOI: 10.1002/mrm.25866
6. Serov V. Fourier Series, Fourier Transform and Their Applications to Mathematical Physics / V. Serov. – New York : Springer International Publishing, 2017. – 534 p. DOI: 10.1007/978-3-319-65262-7.
7. Maggioli F. Orthogonalized fourier polynomials for signal approximation and transfer / F. Maggioli, S. Melzi, M. Ovsjanikov et al.] // Computer Graphics Forum. – 2021. – Vol. 40(2). – P. 435–447. <https://doi.org/10.1111/cgfm.142645>
8. Andrianov I. Asymptotic Methods in the Theory of Plates with Mixed Boundary Conditions / I. Andrianov J. Awrejcewicz, V. Danishevskyy, A. Ivankov. – New York: John Wiley & Sons, 2014. – 288 p. DOI: 201410.1002/9781118725184.
9. Olevska Yu. B. Using of fuzzy mathematical models in automated systems for recognition of high molecular substances / Yu. B. Olevska, V. I. Olevskiy, O. V. Olevskiy // Application of Mathematics in Technical and Natural Sciences: 10th International Conference for Promoting the Application of Mathematics in Technical and Natural Sciences – AMiTaNS'18, Albena, 20–25 June: proceedings. – New York: American Institute of Physics, Melville, NY, 2018. – P. 060003-1–060003-9. – (AIP Conference Proceedings, Vol. 2025(1)). <https://doi.org/10.1063/1.5064911>
10. Prots'ko I. O. Efficient computation of the integer DCT-II for compressing images / I. O. Prots'ko, R. D. Kuzminskij, V. M. Teslyuk, // Radio Electronics, Computer Science, Control. – 2019. – No. 2. – P. 151–157. <https://doi.org/10.15588/1607-3274-2019-2-16>
11. Baker J. A., Jr. Padé approximants / J. A. Baker, Jr. and P. Graves-Morris. – New York: Cambridge University Press, 1996. – 746 p. <https://doi.org/10.1017/cbo9780511530074>
12. Technique of Padé-type multidimensional approximations application for solving some problems in mathematical physics / [I. V. Andrianov, V. I. Olevskiy, I. V. Shapka, T. S. Naumenko] // Application of Mathematics in Technical and Natural Sciences: 10th International Conference for Promoting the Application of Mathematics in Technical and Natural Sciences – AMiTaNS'18, Albena, 20–25 June, 2018: proceedings. – New York : American Institute of Physics, Melville, NY, 2018. – P. 040002-1–040002-9. – (AIP Conference Proceedings, Vol. 2025 (1)). DOI: 10.1063/1.5064886
13. Bosuwan N. Inverse Theorem on Row Sequences of Linear Padé-orthogonal Approximation / N. Bosuwan, López Lagomasino G. // Comput. Methods Funct. Theory. – 2015. – Vol. 15. – P. 529–554. <https://doi.org/10.1007/s40315-015-0121-3>
14. Labych Yu. A. Trigonometric Padé approximants for functions with regularly decreasing Fourier coefficients / Yu. A. Labych, A. P. Starovoitov // Sb. Math. – 2009. – Vol. 200(7). – P. 1051–1074. DOI: 10.1070/SM2009v200n07ABEH004027
15. Buslaev V. I. On the existence of compacta of minimal capacity in the theory of rational approximation of multivalued analytic functions / V. I. Buslaev, S. P. Suetin // J. Approx. Theory. – 2016. – Vol. 206. – P. 48–67. DOI: 10.1016/j.jat.2015.08.002
16. Sablonniere P. Padé-Type Approximants for Multivariate Series of Functions / P. Sablonniere // Lecture Notes in Mathematics. – 1984. – Vol. 1071. – P. 238–251. <https://doi.org/10.1007/bfb0099622>
17. Kida S. Padé-type and Padé approximants in several variables / S. Kida // Appl. Numer. Math. – 1989/90. – Vol. 6. – P. 371–391. [https://doi.org/10.1016/0168-9274\(90\)90027-D](https://doi.org/10.1016/0168-9274(90)90027-D)
18. Application of two-dimensional Padé-type approximants for reducing the Gibbs phenomenon / [Yu. B. Olevska, V. I. Olevskiy, I. V. Shapka, and T. S. Naumenko] / Application of Mathematics in Technical and Natural Sciences: 11th International Conference for Promoting the Application of Mathematics in Technical and Natural Sciences – AMiTaNS'19, Albena, 20–25 June, 2019: proceedings. – New York: American Institute of Physics, Melville, NY, 2018. – P. 060014-1–060014-8. – (AIP Conference Proceedings, Vol. 2164). <https://doi.org/10.1063/1.5130816>
19. Daras N. J. The convergence of Padé-type approximants to holomorphic functions of several complex variables / N. J. Daras // Appl. Numer. Math. – 1989/90. – Vol. 6. – P. 341–360. [https://doi.org/10.1016/0168-9274\(90\)90025-B](https://doi.org/10.1016/0168-9274(90)90025-B)
20. TESTIMAGES free collection of digital images for testing [Electronic resource]. – Access mode: <https://testimages.org/>

## THE METHOD OF ASSESSING THE VALUE OF INFORMATION

**Pilkevych I. A.** – Dr. Sc., Professor, Professor of the Department of Computer Information Technologies, Korolov Zhytomyr Military Institute, Zhytomyr, Ukraine.

**Vakaliuk T. A.** – Doctor of Pedagogical Sciences, Professor, Professor of the Department of Software Engineering, Zhytomyr Polytechnic State University, Zhytomyr, Ukraine.

**Boichenko O. S.** – PhD, Head of the research department of the scientific center, Korolov Zhytomyr Military Institute, Zhytomyr, Ukraine.

### ABSTRACT

**Context.** The task of assessing the value of the institution's information as one of the objects of protection of the information security model is considered.

**Objective.** The goal of the work is the creation of a method of assessing the value of information, which takes into account the time of the final aging of information.

**Method.** The results of the analysis of methods for evaluating the value of information showed that modern approaches are conventionally divided into two directions. In the first direction, the value of information is calculated as the amount of information in bytes. In the second direction, the value of information is calculated in monetary terms. It is shown that modern approaches do not take into account the influence of time on the value of information. A method of assessing the value of information is proposed, which takes into account such characteristics as the term of final aging of information, the level of its access restriction, importance, and form of ownership. The value of information is presented as a quantitative measure that determines the degree of its usefulness for the owner. It is proposed to calculate the value of the initial value of information during its creation or acquisition by calculating the normalized weight of the coefficients according to the formula of the arithmetic mean. It was shown that the current value of information has a functional dependence on the time of existence of information and the time of its final aging.

**Results.** The results of the experiment confirm that the value of information has a nonlinear functional dependence on the time of final aging of information.

**Conclusions.** The conducted experiments confirmed the efficiency of the proposed method of evaluating the value of information and allow recommending it for use in practice to protect the institution's information. Prospects for further research may include the creation of a methodology for assessing the value of an institution's information, taking into account the aging of information and subsequent adjustment of measures to protect it.

**KEYWORDS:** aging of information, the importance of information, restriction of access to information, value of information.

### ABBREVIATIONS

ISO is an International Organization for Standardization;

TMIP is a technical means of information processing;

VOI is a value of information.

### NOMENCLATURE

$n$  is the number of information characteristics that affect the value of information;

$k_i$  is a coefficient that characterizes the quantitative measure of the impact of the characteristics of information on its value;

$k_1$  is a coefficient that characterizes the influence of the level of restriction of access to information on the value of information;

$k_2$  is a coefficient that characterizes the influence of the time of final aging of information on the value of information;

$k_3$  is a coefficient that characterizes the influence of the importance of information on the value of information;

$k_4$  is a coefficient that characterizes the influence of the form of ownership of information on the value of information;

$k_5$  is a coefficient that characterizes the influence of the method of storing information on the value of information;

$L$  is the number of levels of restriction of access to information in the institution;

$l_i$  is a quantitative assessment of the level of restriction of access to information in the institution;

$I$  the number of levels of importance of the institution's information

$i_j$  is a quantitative assessment of the appropriate level of importance of information of the institution;

$F$  the number of forms of ownership of information in the institution;

$f_j$  is a quantitative assessment of the appropriate form of ownership of information of the institution;

$S$  the number of methods of storing information used in the institution;

$s_j$  quantitative assessment of the method of information storage;

$VOI_0$  the initial value of the information of the institution (at the time of its occurrence or receipt);

$t$  is a time from the moment of information occurrence to the moment of determining its value;

$t_1$  is a time from the moment of information occurrence to the moment of its final aging.

## INTRODUCTION

In modern society, the role of information and information resources has significantly increased in all spheres of human life. The transformation of information into a product that has a certain value and corresponding value has led to the emergence of a new object of security – information and information resources. Previously, information security consisted of protecting information and information resources from unauthorized actions. At present, there is a need to protect people, society, or the state from threats that may pose information and information resources. Thus, today the threats that information and information resources can carry affect such aspects of human life as economic, financial, military, technical, and political. The emergence of such threats leads to the improvement of the conceptual model of information security.

The conceptual model of information security conditionally consists of security objects, a model of threats, a model of violators, and an information security system [1–2]. Recently, institutions have increased the cost of developing their information security model. However, not all institutions separately assess the value of information of the institution as one of the objects of security. It also does not take into account the fact that over time, the information loses its value and its further protection becomes impractical.

**The object of study** is the process of assessing the value of information.

**The subject of study** is the method of assessing the value of information.

Known methods [8–10] do not take into account the aging time of information, which has a direct impact on the value of information.

**The purpose of the work** is to develop a method for assessing the value of information, which takes into account the time of final aging of information.

## 1 PROBLEM STATEMENT

Suppose a set of coefficients  $\langle k_i \rangle$  is specified, which characterize the quantitative measure of the influence of information characteristics on its value and take values from the interval  $[0 \dots 1]$ ,  $i = 1, 2, \dots, N$ .

For a given set of coefficients  $\langle k_i \rangle$ , the problem of evaluating the value of information can be represented as the problem of calculating the normalized weight of the coefficients according to the formula of the arithmetic mean.

## 2 REVIEW OF THE LITERATURE

The authors in the scientific works [3–7] reflected on the results of the analysis of the implementation of information security in institutions and proposed to use the international standard of information security ISO/IEC 27001 to manage information-related risks.

In the scientific paper [8] the authors defined the multiple criteria value of information and demonstrated the

potential application when conservation issues conflict with monetary issues.

In the scientific work [9] the authors proposed a new approach to assessing the value of information based on the theory of pattern recognition, which expands its scope and can be successfully implemented using declarative programming languages or universal modeling languages.

An approach to calculating the value of information obtained from the example of a fuzzy mathematical model of the queuing system is presented in a scientific paper [10].

In scientific works [8–10] assessment of the value of information is realized using the methods of information theory, where the main property of information is its quantity.

The method of assessing the information potential with the use of coefficient, cost, and effective methods makes it possible to identify reserves of rational information support of enterprises, to determine the cost of information resources and the effectiveness of their use. This technique is proposed in a scientific paper [11] and can be used by the information security service of institutions to determine the cost of information security of the institution.

The authors in the scientific work [12] proposed to use the coefficient of permissible change, which characterizes the losses of the institution that will not lead to its bankruptcy. This ratio is characterized by the ratio of the growth/decline of the capital of the institution to the growth/decline of the amount of damage from the implementation of information threats.

In the results of the research, which reflected in the scientific work [13], the authors proposed a model for assessing the level of protection of information in the social network from external influences on the information social resource. The result is an assessment of the economic feasibility of implementing an appropriate mechanism of technical means of information protection in social networks, depending on the value of information.

Therefore, today there are two known approaches to assessing the value of information: calculating the amount of information in bytes and calculating the value of information in monetary terms. The first approach makes it possible to protect information without taking into account the level of restriction of access, the level of importance, or the form of ownership of information. The second approach implements the protection of individuals, institutions, and the state from damage that may occur in the event of unauthorized actions with information. The second approach also ensures the integrity, confidentiality, and accessibility of information.

The above approaches do not take into account the influence of time on the relevant properties of information. It is a known fact that the time of existence of information affects its value.

### 3 MATERIALS AND METHODS

The value of information in this study should be understood as a quantitative measure that determines the degree of its usefulness to the owner of the information. The functional dependence of the value of information on its characteristics is reflected in the expression [14–17]:

$$VOI = \frac{\sum_{i=1}^n k_i}{n}. \quad (1)$$

Assessing the value of information based on the methods and techniques of the modern theory of systems analysis provides tools for determining the appropriate coefficients based on the characteristics of information. The coefficients take values from the segment [0 ... 1].

The following characteristics of the information are chosen to assess the value of the institution's information.

#### 1. Level of restriction of access to information.

Restricted information, confidential and/or public information may be processed in the institution. The impact of the level of restriction of access to information on its value is determined by the coefficient  $k_1$ :

$$k_1 = \frac{l_i}{L}, \quad (2)$$

where  $i = [1...L]$ .

The level of restriction of access to information is determined by documents on the organization of information security.

#### 2. Period of final aging of information.

The time of final aging of information depends on the level of access restriction and cannot be longer than the time during which the level of access restriction exists. The impact of the time of final aging of information on its value is determined by the coefficient  $k_2$ :

$$k_2 = \frac{Ol_i}{OL}, \quad (3)$$

where  $i = [1...OL]$ .

#### 3. Importance of information.

The importance of information of the institution should be understood as such information, the loss or unauthorized access to which will cause great damage to the institution or completely stop its work [17]. It is advisable to have several levels of importance of information in the institution. In this case, all information in an institution must have its level of importance (rank). A Group of experts determines the level of importance of information.

The impact of the importance of information on its value is determined by the coefficient  $k_3$ :

$$k_3 = \frac{i_j}{I}. \quad (4)$$

where  $j = [1...I]$ .

#### 4. Form of ownership of information.

The information circulating in the institution may have the following forms of ownership: private, collective, and public. The form of ownership of information must have its level of importance in terms of information security. It is expedient to introduce in the institution an additional division of the collective form of ownership of information into collective information of the institution and collective information of structural units of the institution. The impact of the form of ownership of information on its value is determined by the coefficient  $k_4$ :

$$k_4 = \frac{f_j}{F}, \quad (5)$$

where  $j = [1...F]$ .

A Group of experts determines the form of ownership of information.

#### 5. Method of storing information.

The method of storing information determines on which media it is stored. Depending on the level of restriction of access to information, its importance is determined by the method of storing information. A Group of experts determines the method of storing information.

The impact of the method of storing information on its value is determined by the coefficient  $k_5$ :

$$k_5 = \frac{s_j}{S}, \quad (6)$$

where  $j = [1...S]$ .

The value of information changes over time. As a rule, the value of information decreases over time. The dependence of the value of information on time is determined [14–16]:

$$VOI(t) = VOI_0 \cdot 10^{-\frac{t}{t_1}}. \quad (7)$$

### 4 EXPERIMENTS

An example is considered to verify the method of assessing the value of information of the institution. The institution processes information with the following levels of access: unclassified, confidential, secret, and top secret. According to expression (2), the values of the coefficients are obtained, which are shown in Table 1.

Table 1 – The value of the coefficient  $k_1$

Level of restriction of access to information	Rank	Value
Unclassified	4	0.25
Confidential	3	0.33
Secret	2	0.5
Top secret	1	1

The time of final aging of information according to the appropriate levels of access is set in the institution. According to expression (3), the values of the coefficients are obtained, which are shown in Table 2.

Table 2 – The value of the coefficient  $k_2$

Level of restriction of access to information	Period of final aging	Rank	Value
Unclassified	362	4	0.25
Confidential	1086	3	0.33
Secret	1810	2	0.5
Top secret	3620	1	1

The importance of information in the institution is classified according to the following levels: insignificant, useful, important, and very important [17]. According to expression (4), the values of the coefficients obtained, are shown in Table 3.

Table 3 – The value of the coefficient  $k_3$

Level of restriction of access to information	Period of final aging	Rank	Value
Unclassified	362	4	0.25
Confidential	1086	3	0.33
Secret	1810	2	0.5
Top secret	3620	1	1

The following forms of ownership of information are established in the institution: personal, department of the institution, institution, and state [18]. According to expression (5), the values of the coefficients are obtained, which are shown in Table 4.

Table 4 – The value of the coefficient  $k_4$

Level of restriction of access to information	Rank	Value
Personal	4	0.25
Department of the institution	3	0.33
Institution	2	0.5
State	1	1

The following technical solutions are used to store information in the institution: server, technical means of information processing, and portable storage. According to expression (6), the values of the coefficients are obtained, which are shown in Table 5.

Table 5 – The value of the coefficient  $k_5$

Method of storing information	Rank	Value
Server	3	0.33
Technical means of information processing (TMIP)	2	0.5
Portable storage	1	1

The institution has compiled a list of information in need of protection, which is shown in Table 6.

Substituting the data from tables 1–5 to expression (7) we obtain the initial value of information and the value of information as of 27.01.2022. Calculated values of coefficients that affect the value of information, the initial value of information, and the current value of information are shown in Table 7.

Table 6 – The list information of the institution

Title of the document	Date of create	Period of final aging	Access restriction stamp	Importance	Form of ownership of information	Method of storing information
Activity plan of the department of the institution	05.01.2022	362	Unclassified	Insignificant	Department of the institution	TMIP
A personal plan of the employee of the institution	10.01.2022	362	Unclassified	Insignificant	Personal	TMIP
Enterprise development strategy	05.01.2020	1810	Secret	Important	Institution	Portable storage
Industry development plan	09.01.2021	3620	Top secret	Very important	State	Portable storage
Report on the results of department 1 of the institution for 2021	28.12.2021	1086	Confidential	Useful	Department of the institution	Server
Report on the results of department 2 of the institution for 2021	28.12.2021	3620	Secret	Important	Department of the institution	TMIP
Report on the results of department 1 of the institution for 2020	29.12.2020	1086	Confidential	Useful	Department of the institution	server
Report on the results of department 1 of the institution for 2019	28.12.2019	1086	Confidential	Useful	Department of the institution	Server

Table 7 – Initial and current value of information

Title of the document	$k_1$	$k_2$	$k_3$	$k_4$	$k_5$	Initial value	Current value
Activity plan of the department of the institution	0.1	0.25	0.25	0.33	0.5	0.29	0.25
A personal plan of the employee of the institution	0.1	0.25	0.25	0.25	0.5	0.27	0.24
Enterprise development strategy	0.5	0.5	0.5	0.5	1	0.6	0.23
Industry development plan	1	1	1	1	1	1	0.78
Report on the results of department 1 of the institution for 2021	0.33	0.33	0.33	0.33	0.33	0.33	0.31
Report on the results of department 2 of the institution for 2021	1	0.5	0.5	0.5	0.5	0.5	0.56
Report on the results of department 1 of the institution for 2020	0.33	0.33	0.33	0.33	0.33	0.33	0.14
Report on the results of department 1 of the institution for 2019	0.33	0.33	0.33	0.33	0.33	0.33	0.065

## 5 RESULTS

The study yielded the following results:

1. Assessment of the initial value of information carried out by calculating the average value of the sum of the relevant coefficients. Each of the selected coefficients was calculated using the ranking method. The rank of the relevant information characteristics is determined from the guidance documents on the organization of information security or a specially created group of experts.

The assessment of the current value of information is carried out by taking into account the time of final aging of information and the date of creation (receipt) of relevant information. The value of the current value of information is calculated as the product of the initial value of information by the power factor, which characterizes the aging process of information.

2. In the method of estimating the value of information, a coefficient is calculated that takes into account the place of storage of relevant information. The need to introduce this factor was to take into account additional organizational measures aimed at restricting access to information. Thus, to obtain information stored on removable media, an attacker must obtain it from the information protection service. This approach introduces additional controls and limits the list of people who can carry out an insider threat.

3. Each institution has a list of information that needs protection. According to the level of restriction of access to information, the time of final aging of information is set, as well as its form of ownership and level of importance for the owner. For information from this list, its initial value is calculated.

4. The results of the experiment showed that for information that has the same values as the initial value, over time, the value of the current value of the information decreases. For the information contained in the documents of the first department of the institution, namely in the activity reports for the year, the initial value of the information is 0.33. A year later, the value of information decreased almost 2.4 times, and after 2 years almost 5 times. The results of the experiment confirm that the value of information has a nonlinear functional dependence on the time of final aging of information.

## 6 DISCUSSION

The results of data analysis in Table 7 show that the initial values of information value for a typical document of one department of the institution with identical details, except for the date of creation, are the same. Also, the initial value of information has the same values. At the same time, the current value of the value of information decreases as the time of existence of this information increases. This is explained by the fact that when the time of existence of information approaches the time of final aging of information, the importance of this information is lost.

Taking into account the time of final aging of information for different levels of access to information, (Table 3) will make it possible to evaluate the value of in-

formation and provide information to the head of the institution to make a decision on the feasibility of further expenses for the protection of relevant information.

The method of assessing the value of information consists in finding the initial value of the value of information by calculating the average arithmetic value of the coefficients characterizing the quantitative measure of the influence of the characteristics of the information on its value. When the number of such coefficients increases, the adequacy of the model, which is used to calculate the value of information, will also increase.

## CONCLUSIONS

**The scientific novelty** of obtained results is that the method of assessing the value of information was improved. This method allows obtaining a quantitative value of information, taking into account the level of restriction of access to information, the level of importance of information, period of final aging, method, and place of storage, as well as the form of ownership of information. The proposed method provides an opportunity to automate the process of assessing the value of information on the current date using the mathematical apparatus of the modern theory of systems analysis.

The mathematical model used in the method of assessing the value of information provides an opportunity to investigate the value of information that belongs to the person of the institution and the staff of the institution.

**The practical orientation of the study** is to use the developed method in the information security service of the institution to assess the value of information of the institution and in deciding on the choice of an adequate method of protection of relevant information.

**Prospects for further research** are to study the impact of the proposed set of coefficients based on the characteristics of information for a broad class of practical problems in information security.

## ACKNOWLEDGEMENTS

The author expresses gratitude to Ruslan Hryshchuk, Doctor of technical science, professor for research support and a fruitful paper discussion.

## REFERENCES

1. Pevnev V., Tsuranov M., Zemlianko H., Amelina O. Conceptual Model of Information Security, *Integrated Computer Technologies in Mechanical Engineering*, 2020, Vol. № 188, pp. 158–168. DOI: 10.1007/978-3-030-66717-7\_14.
2. Onyshchenko S., Yanko A., Hlushko A., Sivitska S. Conceptual Principles of Providing the Information Security of the National Economy of Ukraine in the Conditions of Digitalization, *International Journal of Management*, 2020, № 11(12), pp. 1709–1726. DOI: 10.34218/IJM.11.12.2020.157.
3. Hasan Shaikha, Ali Mazen, Kurnia Sherah, Thurasamy Ramayah Evaluating the cyber security readiness of organizations and its influence on performance, *Journal of Information Security and Applications*, 2021, Vol. 58, P. 102726. DOI:10.1016/j.jisa.2020.102726.

4. Palko D., Myrutenko L., Babenko T., Big-dan A. Model of Information Security Critical Incident Risk Assessment, *2020 IEEE International Conference on Problems of Information Science and Technology (PIC S&T)*, 2020, pp. 157–161. DOI: 10.1109/PICST51311.2020.9468107.
5. Fazlida M. R., Said Jamaliah Information Security: Risk, Governance and Implementation Setback, *Procedia Economics and Finance*, 2015, Vol. 28, pp. 243–248. DOI: doi.org/10.1016/S2212-5671(15)01106-5.
6. Mirtsch Mona, Blind Knut, Koch Claudia, Dudek Gabriele Information security management in ICT and non-ICT sector companies: A preventive innovation perspective, *Computers & Security*, 2021, Vol. 109, P. 102383. DOI: doi.org/10.1016/j.cose.2021.102383.
7. Aven T. Risk assessment and risk management: Review of recent advances on their foundation, *European Journal of Operational Research*, 2016, Vol. 253, Issue 1, pp. 1–13. DOI: doi.org/10.1016/j.ejor.2015.12.023.
8. Eyvindson K., Hakanen J., Mönkkönen M., Juutinen A., Karvanen J. Value of information in multiple criteria decision making: an application to forest conservation, *Stochastic Environmental Research and Risk Assessment*, 2019, № 33, pp. 2007–2018. DOI: 10.1007/s00477-019-01745-4.
9. Zaiats V. M. and Zaiats M. M. The figurative approach to calculate the amount of information and estimates its values, *Visnyk Natsionalnoho universytetu "Lvivska politekhnika"*, 2017, 872, pp. 93–100. (Serie: Informatsiini systemy ta merezhi).
10. Zaiats V. M., Rybyska O. M., Zaiats M. M. An approach to evaluating the values and quantity of information in queueing systems based on pattern recognition and fuzzy sets theories, *Kibernetika ta sistemnij analiz*, 2019, Vol. 55, № 4, pp. 133–144.
11. Pererva P.G. Informational activity of the enterprise: management, price and marketing composition, *Bulletin of the National Technical University "KhPI" (economic sciences)*, 2018, № 37 (1313), pp. 120–125.
12. Mokhor V., Davydiuk A. Approach of the information properties destruction risks assessing based on the color scale, *Information Technology and Security*, 2020, Volume 8, Issue 2, pp. 216–223. DOI: doi.org/10.20535/2411-1031.2020.8.2.222608.
13. Laptiev O., V. Sobchuk, A. Sobchuk, S. Laptiev, T. Laptieva Improved model of estimating economic expenditures on the information protection system in social networks, *Electronic Professional Scientific Edition "Cybersecurity: Education, Science, Technique"*, 2020, № 4(12), pp. 19–28. DOI: 10.28925/2663-4023.2021.12.1928.
14. Sawatnatee A., Prakancharoen S. Insider Threat Detection and Prevention Protocol: ITDP, *International Journal of Online and Biomedical Engineering*, 2021, Vol. 17, № 02, pp. 69–89. DOI: doi.org/10.3991/ijoe.v17i02.18297
15. Hmelevskoy R. Research on information security threat assessment of information activity objects, *Modern Information Security*, 2016, № 4, pp. 65–70.
16. Gulak G. M. Metodolohiia zakhystu informatsii. Aspekty kiberbezpeky: pidruchnyk, Kyiv, Vydavnytstvo NA SB Ukrainy, 2020, P. 256.
17. Korchenko O. H., Arkhypov O. Ye., Dreis Yu. O. Otsiniuvannia shkody natsionalnoi bezpetsi Ukrainy u razi vytku derzhavnoi taiemnytsi: Monohrafiia. Kyiv, Nauk.-vyd. tsentr NA SB Ukrainy, 2014, P. 332.
18. Horne C. A., Maynard S. B., Ahmad A. Information security strategy in organisations: review, discussion and future research, *Australasian Journal of Information Systems*, 2014, Vol. 21. DOI: doi.org/10.3127/ajis.v21i0.1427

Received 22.11.2022.  
Accepted 15.12.2022.

УДК 004.93

### МЕТОД ОЦІНЮВАННЯ ЦІННОСТІ ІНФОРМАЦІЇ

**Пількевич І. А.** – д-р техн. наук, професор, професор кафедри комп'ютерних інформаційних технологій Житомирського військового інституту імені С. П. Корольова, Житомир, Україна.

**Вакалюк Т. А.** – д-р педагогічних наук, професор, професор кафедри інженерії програмного забезпечення Державного університету "Житомирська політехніка", Житомир, Україна.

**Бойченко О. С.** – канд. техн. наук, начальник науково-дослідного відділу наукового центру Житомирського військового інституту імені С. П. Корольова, Житомир, Україна.

### АНОТАЦІЯ

**Актуальність.** Розглянуто задачу оцінювання цінності інформації установи, як одного з об'єктів захисту моделі інформаційної безпеки.

**Мета роботи** – створення методу оцінки цінності інформації, що враховує час остаточного старіння інформації.

**Метод.** Результати аналізу методів оцінювання цінності інформації показали, що сучасні підходи умовно поділяються на два напрямки. У першому напрямку цінність інформації обчислюється як кількість інформації в байтах. У другому напрямку цінність інформації обчислюється в грошовому еквіваленті. Показано, що сучасні підходи не враховують вплив часу на цінність інформації. Запропоновано метод оцінки цінності інформації, який враховує такі характеристики, як термін остаточного старіння інформації, рівень її обмеження доступу, важливість і форма власності. Цінність інформації представлена як кількісна міра, яка визначає ступінь її корисності для власника. Пропонується розраховувати величину початкової вартості інформації під час її створення чи отримання шляхом розрахунку нормованої ваги коефіцієнтів за формулою середнього арифметичного. Показано, що поточна цінність інформації має функціональну залежність від часу існування інформації та часу її остаточного старіння.

**Результати.** Результати експерименту підтверджують, що цінність інформації має нелінійну функціональну залежність від часу остаточного старіння інформації.

**Висновки.** Проведені експерименти підтвердили працездатність запропонованого методу оцінювання цінності інформації та дозволяють рекомендувати його для використання на практиці для захисту інформації установи. Перспективи подальших досліджень можуть включати створення методології оцінювання цінності інформації установи з урахуванням часу старіння інформації та подальшим коректуванням заходів із її захисту.



**КЛЮЧОВІ СЛОВА:** старіння інформації, важливість інформації, обмеження доступу до інформації, цінність інформації.

#### ЛІТЕРАТУРА

1. Conceptual Model of Information Security / [V. Pevnev, M. Tsuranov, H. Zemlianko, O. Amelina] // *Integrated Computer Technologies in Mechanical Engineering* – 2020. – Vol. № 188. – P. 158–168. DOI: 10.1007/978-3-030-66717-7\_14.
2. Conceptual Principles of Providing the Information Security of the National Economy of Ukraine in the Conditions of Digitalization / [Onyshchenko S., Yanko A., Hlushko A., Sivitska S.] // *International Journal of Management* – 2020. – № 11(12). – P. 1709–1726. DOI: 10.34218/IJM.11.12.2020.157.
3. Evaluating the cyber security readiness of organizations and its influence on performance / [Shaikha Hasan, Mazen Ali, Sherah Kurnia, Ramayah Thurasamy] // *Journal of Information Security and Applications*. – 2021. – Vol. 58. – P. 102726. DOI:10.1016/j.jisa.2020.102726.
4. Model of Information Security Critical Incident Risk Assessment / [D. Palko, L. Myrutenko, T. Babenko, A. Bigdan] // *2020 IEEE International Conference on Problems of Informatics, Science and Technology (PIC S&T)*, 2020. – P. 157–161. DOI: 10.1109/PICST51311.2020.9468107.
5. Fazlida M. R. Information Security: Risk, Governance and Implementation Setback / M. R. Fazlida, Jamaliah Said // *Procedia Economics and Finance*. – 2015. – Vol. 28. – P. 243–248. DOI: doi.org/10.1016/S2212-5671(15)01106-5.
6. Information security management in ICT and non-ICT sector companies: A preventive innovation perspective / [Mona Mirtsch, Knut Blind, Claudia Koch, Gabriele Dudek] // *Computers & Security*. – 2021. – Vol. 109. – P. 102383. DOI: doi.org/10.1016/j.cose.2021.102383.
7. Aven T. Risk assessment and risk management: Review of recent advances on their foundation / T. Aven // *European Journal of Operational Research*. – 2016. – Vol. 253, Issue 1. – P. 1–13. DOI: doi.org/10.1016/j.ejor.2015.12.023.
8. Value of information in multiple criteria decision making: an application to forest conservation / [K. Eyvindson, J. Hakanen, M. Mönkkönen et al.] // *Stochastic Environmental Research and Risk Assessment*. – 2019. – № 33. – P. 2007–2018. DOI: 10.1007/s00477-019-01745-4.
9. Заяць В. М. Образний підхід до обчислення кількості інформації та оцінки її цінності / В. М. Заяць, М. М. Заяць // *Вісник Національного університету “Львівська політехніка”*, 2017– № 872. – С. 93–100. – (Серія: Інформаційні системи та мережі).
10. Заяць В. М. Підхід до оцінювання цінності та кількості інформації в системах масового обслуговування на основі теорії розпізнавання образів та нечітких множин / В. М. Заяць, О. М. Рибицька, М. М. Заяць // *Кибернетика и системный анализ*. – 2019. – Том 55, № 4. – С. 133–144.
11. Перерва П. Г. Інформаційна діяльність підприємства: управлінська, цінова та маркетингова складові / П. Г. Перерва // *Вісник НТУ «ХПІ»*. Серія: Економічні науки. – 2018. – № 37 (1313). – С. 120–125.
12. Мохор В. Спосіб оцінювання ризиків порушення властивостей інформації за колірною шкалою / В. Мохор, А. Давидюк // *Information Technology and Security*. – 2020. – Vol. 8, Issue 2. – P. 216–223. DOI: doi.org/10.20535/2411-1031.2020.8.2.222608.
13. Удосконалена модель оцінювання економічних витрат на систему захисту інформації в соціальних мережах / [О. А. Лаптев, В. В. Собчук, А. В. Собчук та ін.] // *Кибербезпека: освіта, наука і техніка*. – 2020. – № 4(12). – С. 19–28. DOI: 10.28925/2663-4023.2021.12.1928.
14. Sawatnatee A. Insider Threat Detection and Prevention Protocol: ITDP / A. Sawatnatee, S. Prakancharoen // *International Journal of Online and Biomedical Engineering*. – 2021. – Vol. 17, № 02. – P. 69–89. DOI: doi.org/10.3991/ijoe.v17i02.18297
15. Хмелевський Р. М. Дослідження оцінки загроз інформаційній безпеці об'єктів інформаційної діяльності / Р. М. Хмелевський // *Сучасний захист інформації*. – 2016. – № 4. – С. 65–70.
16. Гулак Г. М. Методологія захисту інформації. Аспекти кібербезпеки : підручник / Г. М. Гулак. – К. : Видавництво НА СБ України, 2020. – 256 с.
17. Корченко О. Г. Оцінювання шкоди національній безпеці України у разі витоку державної таємниці : монографія / О. Г. Корченко, О. Є. Архипов, Ю. О. Дрейс. – К. : Наук.-вид. центр НА СБ України, 2014. – 332 с.
18. Horne C. A. Information security strategy in organisations: review, discussion and future research / C. A. Horne, S. B. Maynard, A. Ahmad // *Australasian Journal of Information Systems*. – 2014. – Vol. 21. DOI: doi.org/10.3127/ajis.v21i0.1427

## ТЕХНОЛОГІЯ ВИПРАВЛЕННЯ ГРАМАТИЧНИХ ПОМИЛОК В УКРАЇНОМОВНОМУ ТЕКСТОВОМУ КОНТЕНТІ НА ОСНОВІ МЕТОДІВ МАШИННОГО НАВЧАННЯ

**Холодна Н. М.** – магістр кафедри «Інформаційні системи та мережі», Національний університет «Львівська політехніка», Львів, Україна.

**Висоцька В. А.** – канд. техн. наук, доцент, доцент кафедри «Інформаційні системи та мережі», Національний університет «Львівська політехніка», Львів, Україна.

### АНОТАЦІЯ

**Актуальність.** Більшість досліджень у напрямі виправлення граматичних та стилістичних помилок зосереджені на корекції помилок в англійськомовному текстовому контенті. Завдяки наявності великих наборів даних досягнуто суттєвого підвищення точності корекції граматики англійської мови. На жаль, досліджень інших мов мало. Системи в для англійської мови постійно розвиваються і наразі активно використовують методи машинного навчання: класифікацію (sequence tagging) та машинний переклад. Для створення якісної моделі машинного навчання для корекції граматичних/стилістичних помилок у текстах тих мов, які є складними морфологічно, необхідна велика кількість паралельних або вручну розмічених даних. Ручна анотація даних вимагає багато зусиль професійних лінгвістів, що робить створення корпусів текстів, особливо морфологічно багатих мов, зокрема, української, часо- та ресурсозатратним процесом.

**Метою дослідження** є розроблення технології виправлення помилок в українськомовних текстах на основі методів машинного навчання з використанням невеликого набору анотованих паралельних даних.

**Метод.** Для даного дослідження при розробці системи корекції помилок в українськомовних текстах із застосуванням оптимального конвеєру (pipeline), що включає в себе попереднє опрацювання текстового контенту, вибір та генерування ознак, обрані алгоритми машинного навчання, в умовах наявності невеликих за обсягом корпусів анотованих даних. Застосування нейронних мереж з новою архітектурою, огляд state-of-the-art методів та порівняння різних етапів конвеєру дасть змогу визначити таку їх комбінацію, яка дозволить отримати якісну модель корекції помилок в українськомовних текстах.

**Результати.** Розроблено модель машинного навчання для корекції помилок в українськомовних текстах. Запропоновано універсальну схему розробки системи корекції помилок для різних мов. Відповідно до отриманих результатів, нейронна мережа має здатність виправляти прості речення, написані українською, однак розроблення повноцінної системи вимагатиме застосування перевірки орфографії за допомогою словників і перевірки правил, як простих, так і заснованих на результаті парсингу залежностей або інших ознак. З-поміж трьох моделей, найкращі показники має попередньо навчена модель нейронного перекладу mT5. З метою економії обчислювальних ресурсів можливим також є застосування попередньо навченої нейронної мережі типу BERT, використовуючи її як у якості енкодера, так і декодера. Така нейронна мережа має вдвічі менше параметрів, ніж інші попередньо навчені моделі машинного перекладу, і показує задовільні результати при виправленні граматичних та стилістичних помилок.

**Висновки.** Створена модель показує відмінні результати класифікації на тестових даних. Розраховані метрики якості машинного перекладу дають змогу лише частково порівняти моделі, оскільки більшість слів і словосполучень у початковому та виправленому реченні співпадають. Найкраще значення як BLEU (0.908), так і METEOR (0.956) отримано для mT5, що співпадає із аналізом прикладів, у якому найбільш точні виправлення помилок без зміни початкового значення речення отримані для такої нейронної мережі. M2M100 має більшу оцінку BLEU (0.847), ніж “Ukrainian Roberta” Encoder-Decoder (0.697), однак, суб’єктивно оцінюючи результати виправлення прикладів, M2M100 значно гірше справляється із подібним завданням, ніж дві інші моделі. Для METEOR також M2M100 (0.925) має більшу оцінку, ніж “Ukrainian Roberta” Encoder-Decoder (0.876).

**КЛЮЧОВІ СЛОВА:** NLP, text pre-processing, корекція помилок, виправлення граматичних помилок, машинне навчання, глибинне навчання, аналіз тексту, класифікація тексту, нейронна мережа.

### АБРЕВІАТУРА

БД – база даних;  
ІС – інтелектуальна система;  
ІТ – інформаційна технологія;  
ПЗ – програмне забезпечення;  
ПО – предметна область;  
BERT – bidirectional encoder representations from transformers;  
GEC – grammatical error correction;  
GECS – grammatical error correction system;  
LSTM – long short-term memory;  
MT – machine translation;  
ML – machine learning;  
NLP – natural language processing;  
NN – neural network;

TPP – text pre-processing.

### НОМЕНКЛАТУРА

$S$  – система граматичної корекції;  
 $I$  – множина вхідних даних;  
 $O$  – множина вихідних даних;  
 $R$  – основні правила опрацювання потоку вхідних даних в ІС граматичної корекції;  
 $U$  – параметри опрацювання вхідних даних;  
 $N$  – нейронна мережа;  
 $\alpha$  – оператор скачування вхідних даних;  
 $\beta$  – оператор опрацювання вхідних даних;  
 $\gamma$  – оператор збереження вхідних даних;  
 $\mu$  – TPP-оператор;

$\chi$  – оператор пошуку помилок;  
 $\omega$  – оператор машинного навчання ІС на достовірних текстових даних;  
 $\lambda$  – оператор граматичної корекції тексту;  
 $i_1$  – множина даних ідентифікації;  
 $i_2$  – множина вхідного текстового контенту;  
 $i_3$  – множина шаблонів/правил помилок;  
 $i_4$  – підтвердження правки від автора/користувача;  
 $o_1$  – маркований/тегований текст з помилками;  
 $o_2$  – колекція пропозицій корекції тексту;  
 $o_3$  – множина підтверджених автором правок;  
 $r_1$  – правила алгоритму взаємодії;  
 $r_2$  – NLP-правила;  
 $r_3$  – правила алгоритму нейронної мережі;  
 $r_4$  – правила алгоритму корекції помилок;  
 $u_1$  – множина рівнів доступу;  
 $u_2$  – множина вимог доступу;  
 $u_3$  – множина NLP-вимог;  
 $u_4$  – множина метрик машинного навчання;  
 $u_5$  – множина вимог корекції помилок.

## ВСТУП

GEC – задача ідентифікації та усунення помилок у вхідному тексті. GEC застосовують в різних сферах, включаючи виправлення пошукових запитів і МТ-результатів, ТРР, перевірку правопису в браузері і текстових процесорах тощо. GEC-методи поділяють на категорії: методи, засновані на правилах; методи, засновані на синтаксичному аналізі речень; статистичне моделювання; класичні ML-методи; МТ на основі глибинного навчання.

Перевірка на основі правил використовує набір попередньо визначених шаблонів помилок для відповідності тексту. Всі правила розробляють зазвичай вручну. Текст є помилковим, якщо відповідає одному з правил [1]. Переваги підходу: швидкодія, інтерпретація результатів, можливості ітеративного розвитку ІС. Однак метод має недоліки: складність ІС збільшується в міру появи різних типів помилок, створення правил є ресурсозатратним і вимагає експертних знань з ПО, зокрема лінгвістики. Одночасно для покриття усіх можливих випадків необхідна величезна кількість правил.

При перевірці на основі синтаксису повністю аналізується морфологія та синтаксис тексту. Для цього потрібна лексична БД, морфологічний і синтаксичний аналізатори (парсери). Залежно від граматики мови, синтаксичний парсер визначає синтаксичну структуру кожного речення у вигляді дерева. Якщо повний аналіз не був успішним, тоді текст є помилковим [1]. Недоліком синтаксичного підходу є необхідність розробки додаткових правил для уточнення необхідних виправлень. Ці правила мають покривати усі можливі варіанти помилок.

Для автоматичного отримання правил із великої кількості тексту використовують статистичні моделі, які навчаються на великій кількості речень і можуть призначати ймовірність новій послідовності слів на основі кількості спостережуваних сполучень слів у

навчальному корпусі. Поширені та більш вірогідні послідовності, які часто зустрічаються в корпусі, вважають правильними, тоді як рідкі послідовності можуть містити помилки [2]. Переваги: автоматичне створення правил та відсутність необхідності застосування експертних знань для опрацювання даних або створення ознак. Недоліки: складна інтерпретація результатів, залежність точності моделі від якості навчального набору даних, необхідність застосування великого за обсягом набору даних.

При класифікації модель навчається передбачувати виправлення для кожного слова/тегу, що позначають дію над певним токеном, яку потрібно виконати для виправлення вхідного речення (sequence labelling). Для класифікації використовують складні системи із рекурентними NN або трансформерами, а також класичні ML-методи: наївний Байєсів класифікатор, метод опорних векторів, випадковий ліс тощо. Останні вимагають ручного створення ознак, як-от POS-теги слів у реченні, парсинг залежностей, відмінки слів, головних і другорядних членів речення тощо. Окрім необхідності застосування експертних знань при побудові моделі, головним недоліком є припущення про незалежність помилок у реченні або контекст слова не містить помилок. Цей метод не дає змоги одночасно виправити кілька співзалежних помилок.

Глибинне навчання архітектури рекурентних NN використовують для багатьох NLP-завдань. На відміну від методу класифікації, моделі глибокого навчання не вимагають розробки ознак, оскільки NN можуть досліджувати їх автоматично. Це є великою перевагою, оскільки генерація ознак вимагає експертних знань з лінгвістики. Особливо популярною варіацією рекурентних NN є LSTM. Рекурентні NN використовують для передбачення тегу, що позначає необхідну дію над токеном для виправлення речення. Дана NN-архітектура використовується у GEC-задачі для МТ тексту, що містить помилки, у його правильний варіант. Для створення кращої МТ-моделі використовують велику кількість пар речень (паралельних корпусів), точність системи у цьому випадку буде залежати від якості набору даних. Окрім того, рекурентні NN сприймають токени послідовно, що сповільнює час навчання та передбачення, унеможливує паралельне опрацювання даних. У випадку комерційного застосування ІС час очікування є критичним, тому все більше досліджень наразі спрямовані на застосування іншої NN-архітектури, що називається трансформер.

Трансформер – модель глибинного навчання, яка замінює механізм рекурентності на механізм уваги, що забезпечує контекст для будь-якого положення токена у вхідній послідовності. Ця властивість надає змогу розпаралелювати набагато більше процесів у порівнянні з рекурентними нейронними мережами, і відтак знижує тривалість навчання [3]. Трансформери швидко стали домінуючою архітектурою для NLP [4], випереджаючи такі альтернативи, як згорткові та

рекуррентні NN, особливо для завдань розуміння мови (класифікація, переказ і узагальнення тексту, MT) та її генерації. Архітектура масштабується відповідно до навчальних даних і фіксує особливості тексту на великій відстані. Попереднє навчання моделі дозволяє навчати трансформери на великих відкритих неанотованих корпусах і згодом легко налаштувати їх до конкретних завдань, отримуючи в результаті високу якість системи.

Підвид трансформерів BERT призначений для попереднього навчання глибоких двонаправлених представлень з неанотованого набору даних. В результаті попередньо навчена модель BERT може бути налаштована лише одним додатковим вихідним шаром для різних NLP-задач без істотних модифікацій внутрішньої архітектури [5].

**Метою дослідження** є проектування та створення GEC-системи в текстах українською мовою за допомогою ML-методів з використанням невеликого набору анованих паралельних даних. До задач, які необхідно вирішити для досягнення поставленої мети, належать:

- опис функціональності та вимог проекрованої ІС;
- порівняння state-of-the-art методів для GEC;
- проектування і застосування нейронних мереж з різною архітектурою;
- вибір найбільш оптимальної моделі у TRP-контексті, векторного вкладення або векторизації, вибору та генерування ознак, ML-алгоритму та його параметрів.

Об'єкт дослідження – процеси ідентифікації та корекції граматичних та стилістичних помилок в україномовному текстовому контенті. Предмет дослідження – методи та засоби виправлення граматичних/стилістичних помилок в україномовних текстах із застосуванням оптимального конвеєру (pipeline) на основі TRP, вибору та генерування ознак, алгоритмів машинного навчання, в умовах наявності невеликих за обсягом корпусів анованих даних. Наукова новизна – застосування нейронних мереж з новою архітектурою, огляд state-of-the-art методів та порівняння різних етапів конвеєру (pipeline) дасть змогу визначити таку їх комбінацію, яка дозволить отримати якісну модель виправлення граматичних помилок в україномовних текстах.

## 1 ПОСТАНОВКА ПРОБЛЕМИ

GEC-систему  $S$  подано коротко:

$$S = \langle I, O, R, U, N, \alpha, \beta, \gamma \rangle,$$

де  $I = \{i_1, i_2, i_3, i_4\}$ ,  $O = \{o_1, o_2, o_3\}$ ,  $R = \{r_1, r_2, r_3, r_4\}$ ,  $U = \{u_1, u_2, u_3, u_4, u_5\}$ .

Основними процесами ІС граматичної корекції є «TRP», «Пошук помилок», «Машинне навчання» та «Виправлення граматичних помилок». TRP-процес ІС граматичної корекції опишемо суперпозицією:

$$C_{AU} = \mu^\circ \beta^\circ \alpha, C_{AU} = \mu(\beta(\alpha(i_1, i_2, i_4), r_1, u_1), u_2).$$

Процес «Пошук помилок» ІС граматичної корекції опишемо суперпозицією:  $C_{CU} = \chi^\circ \beta^\circ \alpha$ , тобто

$$C_{CU} = \chi(\beta(\alpha(C_{AU}, i_2, i_3, i_4), r_1, u_3), r_2).$$

Процес машинного навчання на достовірних даних ІС граматичної корекції опишемо суперпозицією:

$$C_{UL} = \omega^\circ \gamma^\circ \beta^\circ \alpha, C_{UL} = \omega(\gamma(\beta(\alpha(C_{CU}, i_2), i_3), u_4), r_3).$$

Процес «Виправлення граматичних помилок» ІС граматичної корекції на основі GEC-методів опишемо суперпозицією:

$$C_{US} = \lambda^\circ \gamma^\circ \beta^\circ \alpha, C_{US} = \lambda(\gamma(\beta(\alpha(C_{US}, i_2), i_4), u_5), r_4).$$

GEC-методи, засновані на правилах, синтаксичному аналізі або статистичному моделюванні, описані у дослідженнях, що спрямовані на побудову системи перевірки та виправлення текстів, написаних данською [7], грецькою [8], латвійською [9], слов'янською [10], пенджабі [11], філіппінською [12] та арабською [13] мовами. Системи GEC для англійської постійно розвиваються і наразі активно використовують ML-методи: класифікацію (sequence tagging) та MT. Для створення якісної ML-моделі для GEC у текстах тих мов, які є складними морфологічно, необхідна велика кількість паралельних або вручну розмічених даних. Проблему отримання анованих даних без ручного маркування частково вирішують за допомогою алгоритмів, що змінюють початковий текст, додаючи у нього «шум» (noise injection). Окрім того, використовують зворотний переклад безпомилкових текстів для того, щоб отримати їх неграматичні відповідники. Таким чином автоматично згенерують певну кількість розмічених корпусів паралельних текстів [14].

## 2 АНАЛІЗ ЛІТЕРАТУРНИХ ДЖЕРЕЛ

Основними аналогами розробленої ІС є два додатки: Grammarly та LanguageTool.

Grammarly – комерційна онлайн-платформа компанії Grammarly Inc., яка перевіряє й виправляє не тільки граматичні помилки, а й пропонує рекомендації щодо чіткості (стилість та зрозумілість), захопливості (словниковий запас та розмаїття) та тону повідомлення (формальність, ввічливість і впевненість) [16–17]. Для покращення перевірки і точності пропозицій платформа використовує ML-алгоритми і методи глибокого навчання [18]. 30 мільйонів людей і 30 тисяч команд використовують Grammarly щодня [18]. Платформа потрапила до рейтингу Time «Time100 Most Influential Companies of 2022» та FastCompany «The 10 most innovative companies in artificial intelligence of 2022», також до списку Forbes «Cloud 100» та The Software Report «Top 100 Software Companies» [18]. Наразі Grammarly підтримує лише англійську мову та її

діалекти: американський, британській, канадський, австралійський [19].

LanguageTool – GEC-програма із відкритим кодом, що у своїй основі використовує перевірку на основі правил. Усього система більше 30 мов, зокрема і українську [20]. LanguageTool розвивається з 2003 року і наразі налічує 5 415 правил для англійської мови, 1 022 – для української [21]. Щодня користувачі перевіряють більше 20 мільйонів текстів за допомогою цієї платформи [22]. Для перевірки правильності написання україномовних текстів доступні правила таких категорій: варваризми (наприклад, приймати участь – брати), великі літери, граматики, логічні помилки (наприклад, неправильна дата), орфографія, оформлення, пунктуація, стиль, типографія [21]. Недоліком LanguageTool для перевірки україномовних текстів є невелика навіть у порівнянні з англійською мовою кількість правил, які не можуть покрити усі можливі граматичні помилки.

Sai Muralidhar Jayanthi, Danish Pruthi, Graham Neubig розробили систему NeuSpell [23] для виправлення орфографічних помилок англійських текстів, що заснована на методах глибинного навчання. Система складається з 10 різних нейронних мереж, які дають змогу фіксувати контекст навколо орфографічних помилок. Для навчання нейронних мереж використані синтетичні навчальні дані, створені за допомогою кількох стратегій додання «шуму». NeuSpell є програмою із відкритим кодом і наразі підтримує лише англійську мову.

Hunspell [24] – система для виправлення орфографічних помилок, створена для мов з багатою морфологією, складеними словами і різним кодуваннями символів. Використовує спеціальний формат словника, який визначає, які стемми та афікси є дійсними в певній мові. Система забезпечує токенизацію, стемінг і перевірку орфографії для майже будь-якої мови чи алфавіту.

В [6] запропоновано підхід перевірки граматичної правильності текстів на основі «мінімального навчання з вчителем» (minimal supervision). Даний метод полягає у застосуванні невеликого анотованого набору даних і додання штучних помилок до корпусу новин, художньої літератури та інших жанрів (усього 18 мільйонів слів). Дослідники розробили класифікатори для кількох поширених типів граматичних помилок: прийменники, відмінок іменника, форма дієслова та узгодження з дієсловом (розподіл на число та рід). Окрім того, вони застосували метод нейронного МТ тексту, що містить помилки, у його правильний варіант. Alla Rozovskaya та Dan Roth зауважили, що МТ-точність є занадто низькою у зв'язку з недостатньою кількістю навчальних даних. Запропонований метод «мінімального навчання з вчителем» збільшує як точність класифікаторів, так і МТ-якість. За допомогою експериментів показали, що МТ-метод показує незадовільні метрики якості (F-score = 10.6) з використанням набору даних відповідних текстів, що

містить приблизно 200 тис. вручну анотованих і виправлених слів. Цю МТ-систему значно перевершує запропонований у [6] ML-підхід (різновидність навчання з учителем, яке використовує немарковані дані для тренування – зазвичай невелику кількість маркованих даних та велику кількість немаркованих).

Дослідники Oleksiy Syvokon, Olena Nahorna презентували набір даних – корпус текстів [15], що професійно анотований для GEC та вільного редагування українською мовою. Дослідники зібрали тексти з помилками (20 715 речень – 328 779 токенів) від різноманітних авторів, у тому числі носіїв мови. Дані охоплюють найрізноманітніші сфери письма, від текстових чатів і есе до офіційного письма. Професійні коректори виправляли та анотували корпус на помилки, пов'язані з вільним мовленням, граматику, пунктуацією та орфографією. За [1, 15], цей корпус можна використовувати для розробки та оцінки систем GEC українською мовою і дослідження морфологічно багатих мов. За [1], найбільш очевидною проблемою, пов'язаною з GEC-системами, є потреба у високоякісних навчальних корпусах, що містять велику кількість навчальних прикладів.

В [25] розроблено морфологічний парсер rymorphu2 для української мови у форматі бібліотеки для мови програмування Python. rymorphu2 аналізує частину мови, число, відмінок, час, лему та стем заданого слова. Морфологічний парсер заснований на словниках OpenCorpora, що конвертовані до формату XML. Користувачі також мають змогу додавати власні слова та правила, це дає можливість проводити морфологічний аналіз текстів певної ПО без зміни вихідного коду rymorphu2 та адаптувати rymorphu2 для роботи з іншими мовами.

В [26] розроблено систему TRP, морфологічного та синтаксичного аналізу україномовних текстів. Для токенизації, розподілу на речення та пошуку email-адрес дослідники використали бібліотеку NLTK мови програмування Python та регулярні вирази, для видалення стоп-слів та пошуку іменованих сутностей – відповідні словники. Для морфологічного аналізу слів, їх лемматизації або стемінгу була використана бібліотека rymorphu2, що підтримує українську мову. Окрім того, дослідники реалізували графічний інтерфейс додатку за допомогою бібліотеки PyQt.

В [27] проаналізовано існуючі методи пошуку іменних груп в англійських і україномовних текстах та розробили метод детектування іменних груп на основі дерева залежностей речення і моделі розпізнавання іменованих сутностей. Довели, що методи аналізу англійського тексту не можуть бути використані для україномовних документів, адже вони створені з урахуванням особливостей структури побудови речень тільки в англійських текстах.

В [28] адаптовано алгоритм стемінгу Портера для TRP україномовних текстів. Окрім того, дослідники реалізували алгоритм на мові програмування PHP і виклали його у вільний доступ. Довели, що кращого

результату можна досягти через лематизацію, однак для цього потрібно провести додаткові дослідження.

В [29] розроблено GEC-систему в англійських текстах за допомогою спеціальної моделі глибинного навчання – Transformer [3]. Виправлення помилок відбувається завдяки використанню даної моделі для задачі перекладу неправильного тексту у його відкоригований відповідник. Особливістю даної системи є вибір TPP-методів: автоматичне виправлення орфографічних помилок, токенизація [30], метод кодування ознак Byte Pair Encoding. Довели, що дана ІС має достатню точність (F-score = 60.93 на завданні CoNLL-2014 [31]) і швидкість (10 слів за секунду) для подальшого її впровадження і застосування на онлайн-платформах.

В [32] застосовано нейронну мережу архітектури encoder-decoder, що складається із шарів згортки та механізму уваги, для GEC у англійських текстах за допомогою МТ. Для процесу декодування використані ансамблі однакових моделей, що ініціалізовані випадковим чином. Для фінального вибору речень з-поміж можливих кандидатів оцінки їх ймовірностей змінені за допомогою додатково створених ознак. ІС значно перевершила якість попередньо існуючих МТ-систем, окрім того і на основі рекурентних нейронних мереж.

В [33] розроблено GEC-систему у англійських текстах, що заснована на підході анотації послідовностей (sequence tagging). Окрім стандартних тегів (keep, delete, append, replace), що позначають дію, яку необхідно виконати над токеном для виправлення речення, також запропоновані так звані g-transformations: зміна регістру першої літери, об'єднання або розділення токена, зміна числа іменника або форми дієслова. Для виправлень токенів застосований ітеративний підхід, у якості ML-методу – різні моделі сімейства Transformer. Такий підхід дозволив отримати найкращі на момент публікації статті оцінки якості системи та збільшення швидкості опрацювання даних у 10 разів у порівнянні із іншими ІС, заснованих на архітектурі Transformer.

В [34] описано GEC-систему, засновану на підході анотації послідовностей за допомогою класичного ML-алгоритму – наївного класифікатора Байєса. Усього використано п'ять моделей – відповідно до наявних у корпусі типів граматичних помилок. Попри те, що система була однією з найкращих на завданні CoNLL [35], аналіз результатів показує, що досягнення дуже високої точності при граматичній корекції вимагає більш складних NLP-методів.

В [9] розроблено засновану на правилах програму перевірки граматики для латвійської мови. Усього реалізовано дві групи правил: правила, що описують правильні речення, і правила, що описують граматичні помилки. На корпусі текстів, написаних людьми, що не є носіями мови, система досягла F1-score 62,4% і 40,2% на корпусі студентських робіт. Недолік: не перевіряє правильність слова в залежності від його контексту. Відповідно, правильно написані та

узгоджені слова, що не належать реченню, будуть маркуватися як безпомилкові.

В [8] описано ІС для перевірки правильності текстів, написаних на грецькій мові. Система заснована на правилах та парсингу синтаксису речення. ІС аналізує текст користувача та надає виправлення, опис помилок, а також правила щодо стилю та семантичної інформації в тексті. Системне оцінювання проводилося як паралельне виправлення одних і тих самих текстів програмою та людиною. Система перевірки грецької граматики наблизилася до 90% виправлення людського. Однак ця ІС не може опрацювати всі можливі граматичні помилки.

В [36] розроблено універсальну ML-модель для GEC у текстах, написаних на різних мовах. Дослідники стверджують, що для отримання якісної багатомовної GEC-системи мають бути виконані такі два кроки: автоматична генерація достатньої кількості навчальних даних і використання ML-моделей сімейства Transformer із величезною кількістю параметрів (до 11 мільярдів). Таким чином, автори досягнули високої точності виправлення помилок для чотирьох мов: англійської, чеської, німецької.

В [37] досліджено здатність GEC-моделей до узагальнення граматичних правил для корекції нових помилок у тексті. ML-модель основі Transformer протестована з використанням невідомих їй прикладів. Отримано незадовільні результати навіть у випадку простих правил і зменшеного словнику, що може свідчити про те, що алгоритму бракує можливості узагальнення, необхідної для виправлення нових помилок у наданих тестових прикладах. Для більш якісного GEC за меншого обсягу навчального корпусу існуючі ІС на основі ML варто поєднати з окремою перевіркою певних правил.

В [38] запропоновано мовно-незалежну стратегію для розробки багатомовної GEC-системи. Для її імплементації необхідні лише попередньо навчена МТ-модель та корпус паралельних навчальних даних для перекладу з англійської на обрану мову. Перш за все, МТ-модель генерує нові синтетичні дані на обраній мові, що застосовують у якості помилкових вхідних текстів. Отриманий корпус використовують для попереднього навчання моделі, після чого її потрібно додатково налаштувати, використовуючи анотований набір даних для GEC у текстах обраної мови. Досить гарні показники точності виправлення досягнуті для німецької, китайської мов.

В [39] адаптовано фреймворк Break-It-Fix-It (BIFI), що оригінально використовувався для виправлення коду програм за умови відсутності ідеальних зразків, до завдання перевірки граматичної правильності речень. Дослідники використали попередньо навчену мовну модель для розробки бінарного класифікатора LM-Critic, який виконує завдання попереднього сортування речень і вказує, чи містить певний текст граматичні помилки. LM-Critic визначає речення граматично правильним, якщо йому відповідає більший числовий показник ймовірності, ніж його

неправильним «сусідам». Утворені пари правильних і неправильних речень використовують для навчання «коректора», що опрацьовує лише речення, позначені LM-Critic як ті, що містять граматичні помилки.

В [40] для вирішення проблеми необхідності застосування великих за обсягом корпусів навчальних даних згенерували помилкові версії великих неанотованих наборів текстів за допомогою запропонованої функції шуму. Отримані паралельні корпуси згодом використовують для попереднього навчання моделей на основі архітектури Transformer. Потім необхідно додатково налаштувати ML-модель відповідно до ПО та стилю датасету. Дана GEC-система ускладнена використанням нейронної перевірки орфографії, що сортує запропоновані варіанти виправлень в залежності від контексту.

В [41] змінили архітектуру ML-моделі Transformer, впровадивши новий механізм уваги, що заснований на дереві залежностей слів у реченні. Так як неправильно побудоване речення може спричинити помилку у парсингу дерева залежностей, дослідники також навчили NN виправляти графи, що містять помилки. Додатково застосувавши запропонований метод аугментації даних, отримали гарні показники GEC-точності навіть без попереднього навчання NN.

В [42] застосували нейронну модель sequence-to-sequence на рівні символів для того, щоб уникнути проблем зі словами OOV (out of vocabulary words, відсутні у словнику). У якості енкодера застосовують двонапряму рекурентну NN, декодер – також рекурентна нейронна мережа, поєднана із механізмом уваги. Декодер генерує вихідне речення посимвольно. Навіть попри те, що такий підхід допомагає нейронній мережі опрацьовувати невідомі їй слова, вона не може ефективно опрацьовувати інформацію на рівні слова: ця модель отримала оцінку F-score = 40,56 на тестовому наборі CoNLL [31].

В [43] застосували новий підхід до автоматичної генерації навчальних анотованих прикладів та GEC за допомогою генеративних змагальних мереж (GANs). Генеративні змагальні мережі складаються з генератора, що постійно генерує навчальні приклади, і дискримінатора, що класифікує отримані дані. У [43] генератором є NN Grammatical Error Labeler, що навчається додавати помилки, аналогічні помилкам у навчальній вибірці текстів, дискримінатор – NN Grammatical Error Detector, що навчається розпізнавати правильну мітку, що позначає дію, яку необхідно виконати над певним токеном.

В [44] розробили дизайн платформи для GEC і створили відповідну ML-модель, поєднану з перевіркою правил. Система складається з трьох модулів: виправлення помилок, адміністрування системи, фільтрування відгуків. GEC-Модуль включає TPR. На TPR-етапі також визначається валідність введених користувачем даних. Модуль фільтрування відгуків дозволяє користувачам скорегувати неправильні виправлення системи. Такий підхід дає змогу виправити можливі помилки, наявні

у наборі даних, оскільки отримані відгуки користувачів можуть використовуватись для додаткового навчання нейронної мережі.

В [45] презентували корпус анотованих даних для створення GEC-системи та редагування текстів, написаних чеською мовою. Усього набір даних містить 42 210 речень. Окрім того, дослідники використали MT-підхід для GEC. Нейронна мережа архітектури Transformer попередньо навчена на додатково синтезованих з головного набору даних реченнях. Окрім чеської мови, автори статті застосували даний підхід для виправлення помилок в німецькомовних текстах.

В [46] запропонували та реалізували новий метод автоматичного створення навчальних даних із застосуванням двох різних за якістю моделей перекладу. «Погана» система перекладу є статистичною моделлю на рівні речень, якісна – навченою нейронною мережею. Даний метод дав змогу отримати 10 мільйонів паралельних речень для навчання нейронної мережі архітектури Transformer. Даний метод дасть змогу отримати великі датасети навчальних даних для інших мов, для яких великі корпуси анотованих даних не є доступними.

В [47] використали генеративні змагальні мережі для створення системи автоматичного виправлення помилок у англійських текстах. На відміну від алгоритму в [43], в даному дослідженні генератор виправляє помилки, «перекладаючи» помилкові речення у правильні, а дискримінатор навчається оцінювати якість перекладу та наявність помилок у опрацьованому генератором реченні. Дискримінатор, що приймає на вхід два речення, заснований на сіамських рекурентних або згорткових NN.

### 3 МАТЕРІАЛИ ТА МЕТОДИ

Система перевірки граматичної правильності речень може використовуватись для перегляду запропонованих змін та їх пояснень, а також для автоматичного виправлення правильності речень.

Така система може використовуватись як і індивідуальним користувачем для перевірки власних текстів, так і сторонньою NLP-програмою для попереднього або фінального опрацювання речень.

Зовнішніми сутностями є: користувач, NLP-програма, адміністратор системи.

Зацікавлені особи прецеденту та їх вимоги:

- користувач створює обліковий запис, завантажує або відкриває документи, перевіряє граматичну правильність речень, застосовує запропоновані зміни;
- стороння NLP-програма відправляє текст на перевірку за допомогою API системи, отримує змінений текст або перелік можливих виправлень;
- адміністратор системи виконує її налаштування, навчання нейронної мережі, перевірку запропонованих користувачем виправлень.

Користувач ПС: 1) фізична особа, що використовує систему для перевірки власних текстів; 2) NLP-програма.

#### Передумови прецеденту:

- комп'ютер, за допомогою якого здійснюватиметься аналіз, підключений до Інтернету та має встановлене необхідне ПЗ;
- користувач має бути успішно авторизованим у системі;
- словник (якщо використовується перевірка орфографії) містить достатню кількість слів і (або) правил їх утворення, є доступним;
- реалізовані усі ключові модулі, що забезпечують основний функціонал системи;
- штучна нейронна мережа (якщо застосовуються ML-методи) є попередньо натренованою на виконання необхідного завдання і має достатньо високі показники якості.

#### Основний успішний сценарій:

- користувач завантажує програму або розширення, або відкриває онлайн-додаток у браузері;
- користувач завантажує або створює документ і заповнює його;
- система виконує перевірку тексту;
- система відображує необхідні виправлення;
- користувач підтверджує або скасовує застосування запропонованих виправлень;
- користувач зберігає документ. Зберігається також і історія змін.

#### Альтернативні потоки:

- Власний файл з даними не відповідає типу файлу, який може відкрити програма.

1. Програма повідомляє користувача про помилку і скасовує відкриття файлу.

2. Користувач обирає файл з правильним розширенням.

3. Програма відкриває файл з правильним розширенням, зчитує дані, проводить їх попередню обробку (точка повернення в основний сценарій).

#### – Помилка у роботі програми:

1. Користувач звертається до служби технічної підтримки (розробника) і повідомляє про помилку у програмі.

2. Розробник усуває помилку і надає нову версію або надає інформацію про способи її самостійного усунення (точка повернення).

#### Пост-умови:

– Користувач запропонував власний ГЕС-приклад у реченні у випадку отриманих незадовільних результатів;

– Запропоновані користувачем виправлення переглянуті адміністратором для подальшого налаштування системи або збережені для персоналізації подальших виправлень;

– Дані, налаштування і файли користувача занесені до бази даних.

#### Спеціальні СВ:

– ІС повинна відповідати бажаному рівню точності, яка перевіряється спочатку на навчальних даних, потім – на тестових даних, у режимі перевірки власного файлу точність не перевіряється;

– система має підсвічувати неправильні слова та словосполучення, надавати обґрунтування запропонованих змін;

– для доступу до функціоналу системи сторонній NLP-додаток використовує її API-сервіс.

#### Список необхідних ІТ та додаткових пристроїв:

– користувач повинен мати комп'ютер із встановленим ПЗ та підключенням до мережі Інтернет для отримання даних;

– система використовує розподілену нереляційну базу даних для зберігання даних користувачів;

– для додаткового опрацювання текстів використовуються попередньо навчені ML-моделі.

Для побудови діаграми варіантів використання для системи автоматичного виправлення помилок україномовних текстів (рис. 1) спочатку необхідно визначити основних акторів: Користувач (User) і NLP-Програма (NLP Application), а також основні варіанти використання як Apply correction (застосувати корекцію), Check text (перевірити текст), Suggest correction (запропонуйте виправлення), Upload text (завантажити текст), Check text for all types of errors (перевірити текст на всі види помилок), View suggestions (переглянути пропозиції).

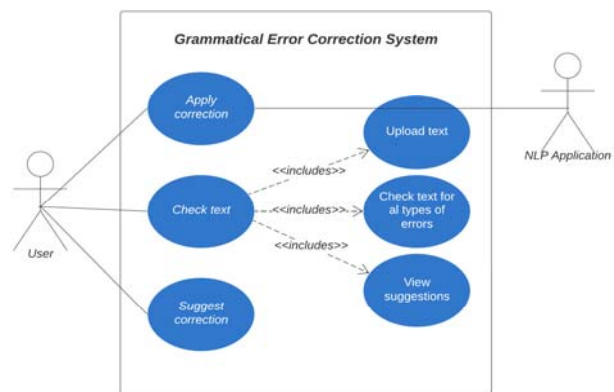


Рисунок 1 – Діаграма варіантів використання GECS

GECS – це десктопний / мобільний додаток, онлайн-платформу або модуль-доповнення до певного ПЗ, наприклад, браузеру або текстового редактору. ПЗ для аналізу природної мови у свою чергу може використовувати API сервісу автоматичного виправлення речень, обмінюючись повідомленнями за допомогою архітектури REST взаємодії між додатками. ГЕС-система включає в себе також сервер опрацювання даних, оскільки за умови застосування ML-методів, у тому числі і мовних моделей на основі архітектури Transformer з мільйонами параметрів, додаток потребуватиме багато обчислювальних ресурсів для обробки текстів. Збільшення кількості користувачів, і, відповідно, збільшення обсягу даних також потребуватимуть додаткових комп'ютерних потужностей. До складу вищезазначеної ІС також входить сервер БД користувачів, що зберігає документи у обліковому записі. Взаємодія із серверами опрацювання та



зберігання даних відбувається через графічний інтерфейс ІС, онлайн-платформи або доповнення. Альтернативним варіантом є встановлення запропонованого ПЗ на потужному комп'ютері, що має достатньо ресурсів для опрацювання даних локально. У іншому випадку, локальна версія ПЗ може мати спрощений функціонал. На рис. 2 подана розширена діаграма варіантів використання із уточненням функціоналу ІС з відповідними варіантами використання як Apply correction (застосувати корекцію), Check text (перевірити текст), Suggest correction (запропонуйте виправлення), Create a document (створити документ), Login/Signup (вхід/вихід), Upload text (завантажити текст), Check text for all types of errors (перевірити текст на всі види помилок), View suggestions (переглянути пропозиції), Upload a document (завантажити документ), Send request (відправляти запит), Get list of suggested corrected (отримати список запропонованих виправлень).

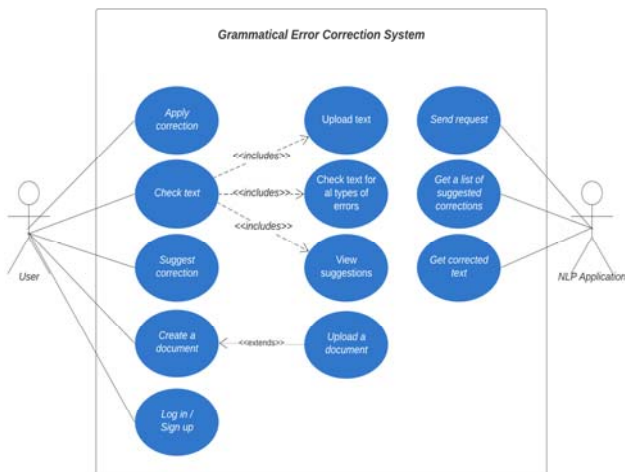


Рисунок 2 – Доповнена діаграма GECS

В залежності від виду взаємодії користувача і GECS не всі типи перевірок можуть бути доступними. Для зберігання та редагування особистих файлів користувач має бути зареєстрованим та авторизованим (варіант використання Log in / Sign up) у системі. Документи, логін, пароль та інші дані користувачів зберігаються на окремому сервері.

Варіант використання «Перевірити текст» (Check text) включає в себе три інших: Завантажити текст, Перевірити текст на всі типи помилок, Переглянути пропозиції. Завантаження тексту може відбуватись у кілька варіантів: вставлення свого тексту у textbox, відкриття файлу або створення нового документу. Створення (завантаження) документу виділено у окремий варіант використання, оскільки даний функціонал може бути реалізований лише для онлайн-платформи або локально встановленого додатку. До базових належать помилки, пов'язані з вільним мовленням, граматиною, пунктуацією та орфографією. В залежності від реалізованого функціоналу система може пропонувати заміни слів

на синоніми для збільшення лексичної різноманітності тексту або заміни відповідно до певного стилю письма. Варіант використання Переглянути пропозиції позначає почерговий перегляд запропонованих виправлень разом із поясненням порушеного правила або причини заміни. Окрім того, частини тексту, що містять помилки, мають бути підсвічені відповідно до типу порушених правил. Користувач повинен мати змогу застосувати лише обрані виправлення, ті, які є необхідними у даному випадку. Якщо користувач вважає якість запропонованих виправлень незадовільною, він повинен мати змогу самостійно виправити текст і зберегти його як зразок для подальшого налаштування системи або навчання нейронної мережі. Таким чином, наступні виправлення системи матимуть кращу якість і будуть більш персоналізованими для кожного окремого користувача. Варіант використання Застосувати виправлення дає змогу користувачеві автоматично виправити усі знайдені системою граматичні помилки без перегляду детальних пояснень порушених правил.

NLP-Додатки також можуть використовувати GEC-систему для попередньої або пост-обробки текстів. В цьому випадку додаток має відправити запит, використовуючи Application Programming Interface (API) сервісу, система – відправити результат аналізу або одразу виправлений текст. Отже, GEC може бути реалізоване як для індивідуального використання у вигляді застосунку, веб-сайту або розширення для додатків, так і як модуль у системі аналізу або генерації природної мови.

Для побудови діаграми класів (рис. 3) у контексті інформаційної GEC-системи визначено вісім основних класів: GEC Software (Програма, що реалізує користувацький інтерфейс системи), User (Користувач), User Database (БД користувачів), Document (Документ), Data Processing Server (Сервер обробки даних), Check Result (Результат перевірки), Dictionary (Словник), ML model (ML-Модель), NLP application (NLP-Додаток).

Клас User (користувач) описує індивідуального користувача системи, який завантажує додаток, встановлює розширення або перевіряє тексти за допомогою онлайн-платформи. Цей клас містить такі атрибути як ім'я, електронна пошта і дані для авторизації у програмі – логін і пароль. Відповідно, він може авторизуватися у системі, завантажити і перевірити текст, створити і зберегти документ, застосувати виправлення, надане системою, або запропонувати власне. Також, у контексті інформаційної системи, користувач створює документ, що зберігає безпосередньо сам текст, а також мета-дані про історію внесених змін і виправлень, автора документу і дату створення. Клас User Database (БД користувача) може зберігати, записувати, оновлювати, повертати інформацію і документи, додати або верифікувати користувача.

Клас GEC Software (ПЗ GEC) постає посередником між користувачем, сервером обробки даних та базою даних. ПЗ може авторизувати користувача, запросити дані із сховища, відправити текст на аналіз, отримати результати та візуалізувати їх, застосувати запропоновані виправлення і зберегти результат, зберегти введені користувачем виправлення. Як вже зазначалось, альтернативним варіантом є використання локальних обчислювальних ресурсів без під'єднання до зовнішнього серверу. GEC Software також надає API сервісу для запитів на аналіз текстів та автоматичне виправлення помилок: до атрибутів цього класу належать тип запиту, адреса для запиту і тип даних, що відправляє сервер як відповідь. Даний ресурс може обробляти запити, автоматично застосовувати необхідні виправлення та відправляти правильний текст на сервер, з якого прийшов запит. Подібні сервіси визначають тип користувача за токеном.

Клас NLP Application (NLP-застосунок) позначає будь-яку програму для аналізу або генерації природної мови, що використовує GEC-систему для попередньої або пост-обробки даних. Така програма має свою структуру і призначення, метод `doltsMagic()` відповідає за пряме її застосування. Ця програма може відправляти запити певного типу і обробляти пакет даних, отриманий від додатку GEC Software.

Клас Data Processing Server (сервер опрацювання даних) описує поведінку серверу обробки даних. Такі сервери мають свої технічні характеристики (кількість ядер процесора, об'єм операційної та

постійної пам'яті, наявність графічного процесора тощо), які напряду впливають на швидкість обробки великого масиву даних.

Для коректного аналізу граматичної коректності речень сервер повинен спочатку виконати TPP, після цього – представити записи як вектори. Потім – парсинг дерева залежностей слів у реченні, частин мови та іменованих сутностей. Отримане дерево залежностей аналізується за допомогою правил. Для автоматичного виправлення використовується підхід передбачення міток-тегів, що вказують на дію, яку необхідно виконати над кожним токеном. Для речень із багатьма помилками, парсинг дерев залежностей яких дає некоректні результати, застосовується MT-метод тексту із помилками у правильний його варіант.

Перевірка орфографії є першим етапом перевірки граматичної коректності речень. Словник Dictionary позначає окремий клас тому, що для перевірки орфографії використовують словники певної мови, що також можуть містити правила утворення слів, а результати ранжуються за зростанням Edit Distance або модифікаціями цієї метрики.

Клас ML model (ML-модель) описує поведінку ML-моделі, до методів якого належать тренування моделі на навчальних даних та передбачення результатів. Кратність зв'язків «0...\*» та відношення агрегації вказують на те, що система не має обов'язково використовувати MT-алгоритми для перевірки граматичної правильності речень та їх автоматичного виправлення.

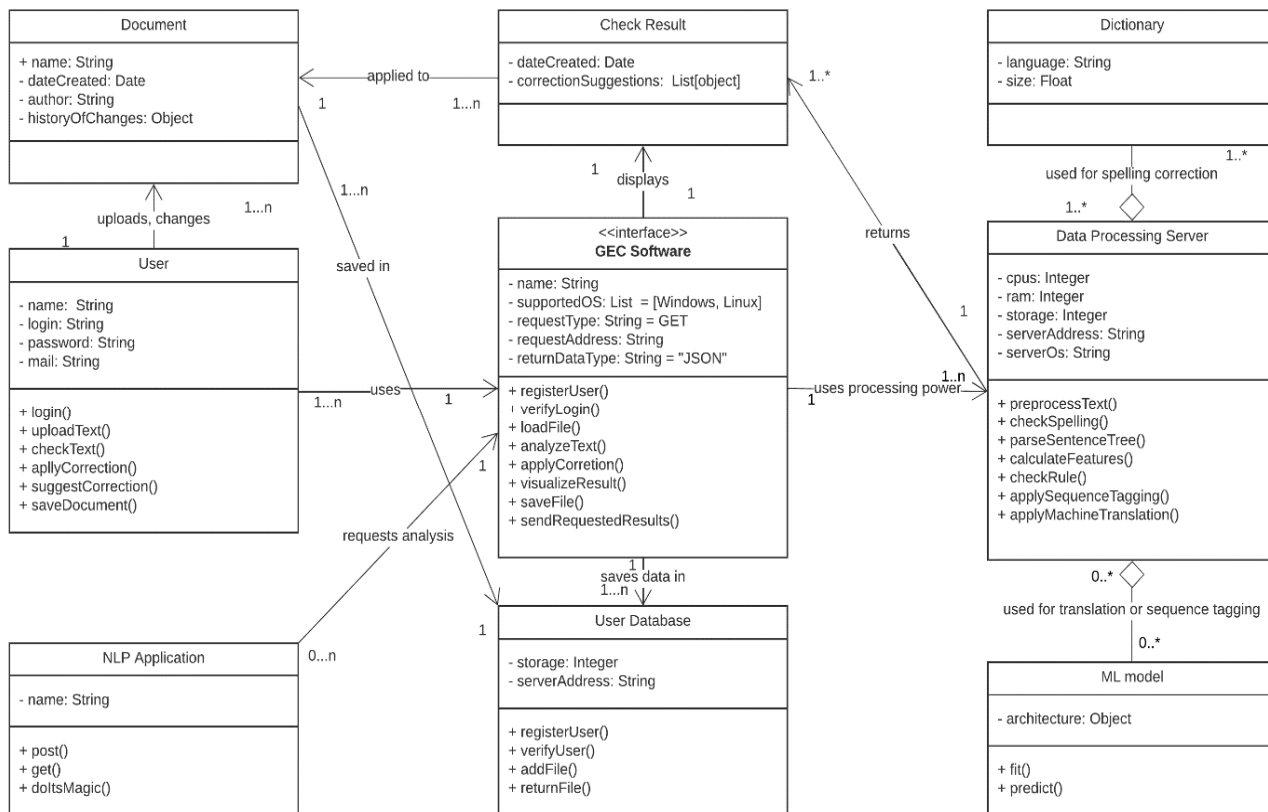


Рисунок 3 – Діаграма класів для GEC-системи

Клас Check Result (результат перевірки) означає результат перевірки одного тексту. Його основним атрибутом є список запропонованих виправлень. Виправлення містять індекси початкового і кінцевого символів, виправлену стрічку, пояснення порушеного правила або причини необхідної заміни. Кратності усіх зв'язків та їх назви вказані на рис. 3.

Послідовність кроків взаємодії із системою перевірки граматичної правильності текстів може бути представлена за допомогою двох діаграм послідовності відповідно для індивідуального користувача додатку та для NLP-застосунку, який використовує виправлення текстів у граматично правильний варіант як одну із складових власного конвеєру (pipeline).

На рис. 4 подано діаграму послідовності для режиму перевірки власного тексту користувача відповідними компонентами як Software (програмне забезпечення), ML-модель Processing server (сервер опрацювання), ML model (), User Database (база даних користувачів). На схемі пропущені кроки реєстрації та авторизації, а також створення і збереження документу. Основний алгоритм такий:

1. Завантажити текст (Load text).
2. Надіслати текст (Send text).
3. Перевірити орфографію (Check spelling).
4. Виправити орфографічні помилки (Correct spelling errors).

5. Проаналізувати дерево залежностей або непередбачених ситуацій (Parse dependency or contingency tree).

6. Перевірити розширені правила на основі дерева (Check advanced rules based on tree).

7. Перевірити за допомогою моделі машинного навчання (Check with machine learning model).

8. Застосувати корекцію машинного навчання (Apply machine learning correction).

9. Повернути результати (Return results).

10. Повернути пропозиції у форматі: [i\_початок, i\_кінець, пропозиція, пояснення] (Return suggestions in format: [i\_start, i\_end, suggestion, explanation]).

11. Показати пропозиції (Display suggestions).

12. Застосувати виправлення (Apply fixes).

13. Запропонувати виправлення (Suggest correction).

14. Запам'ятати вподобання (Remember preferences).

БД користувачів використовується для збереження вподобань / зразків виправлень користувачів для подальшої персоналізації запропонованих системою виправлень. Спочатку користувач має обрати варіант завантаження даних – вставкою тексту, відкриття файлу або створення і заповнення нового документу. ПЗ, встановлене на персональному комп'ютері, або доступне у форматі онлайн-сервісу чи розширення, відправляє дані для аналізу на сервер. На цій діаграмі послідовностей наведений лише один із можливих підходів до побудови системи граматичної

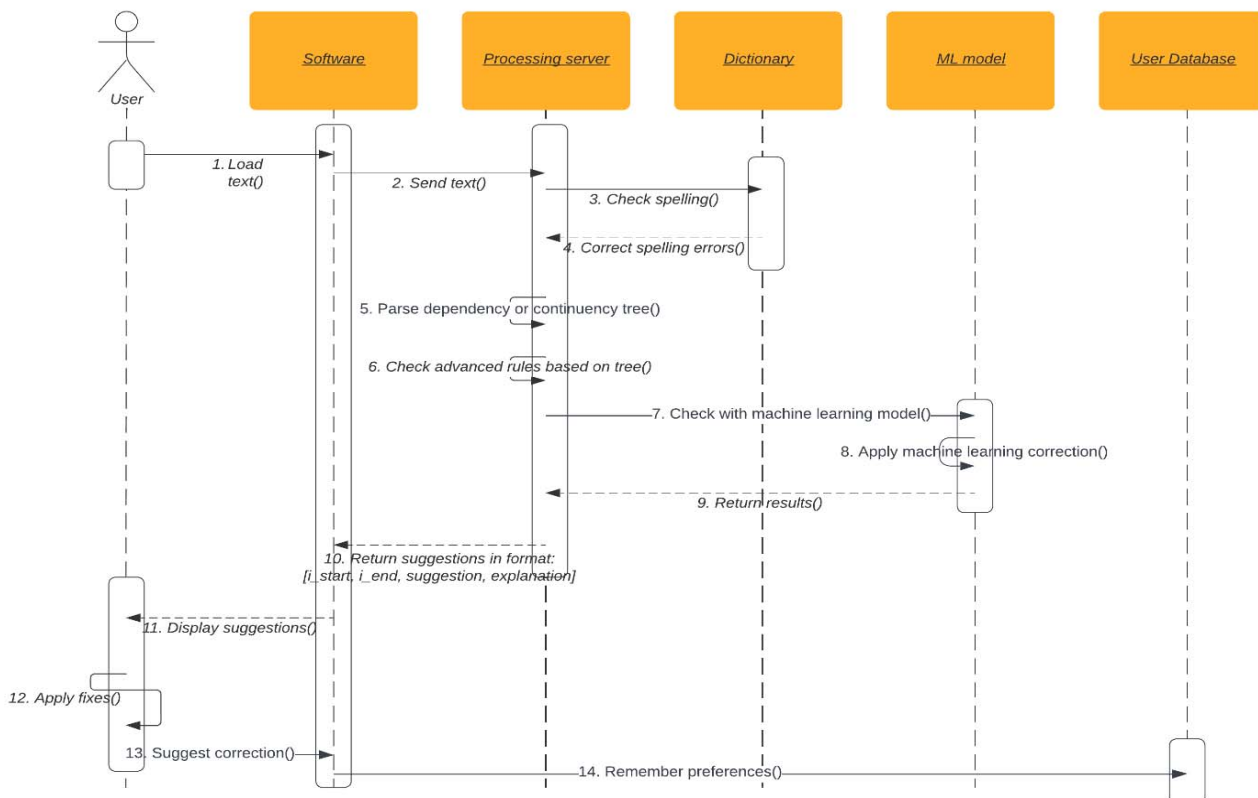


Рисунок 4 – Діаграма послідовності для першого режиму роботи системи

корекції, що поєднує у собі перевірку орфографії, правил, парсинг синтаксичних дерев, перевірку правил на основі дерев і ML-алгоритми. Після перевірки орфографії, що відбувається з використанням словника певної мови, система має перевірити виконання базових правил, що не потребують парсингу синтаксичних дерев, як-от чергування «у-в», правопис «пів-напів», вживання апострофа тощо. Застосування перевірки правил та парсингу синтаксичних дерев обумовлене двома факторами:

– відсутністю достатньо великих анотованих / паралельних корпусів української мови для навчання повністю автоматичних моделей, що засновані на алгоритмах глибинного навчання;

– недостатньо спроможністю мовних моделей штучного інтелекту узагальнювати правила.

У випадку, коли у системі не реалізовані усі можливі правила правопису, в тому числі і ті, що засновані на парсингу синтаксичних дерев, одним із варіантів є додаткове використання ML-методів для передбачення тегів, що позначають дію, яку необхідно виконати над токенами, або для «перекладу» тексту із помилками у правильний варіант. Надалі ПЗ візуалізує і застосовує запропоновані виправлення, зберігає документ у базі даних за необхідності.

Другий режим застосування системи – аналіз граматичної коректності отриманих за допомогою API запитів (рис. 5). На цій діаграмі послідовності зображені етапи відправки, отримання і опрацювання запитів від стороннього NLP додатку (NLP Application) до системи (GEC System), зокрема:

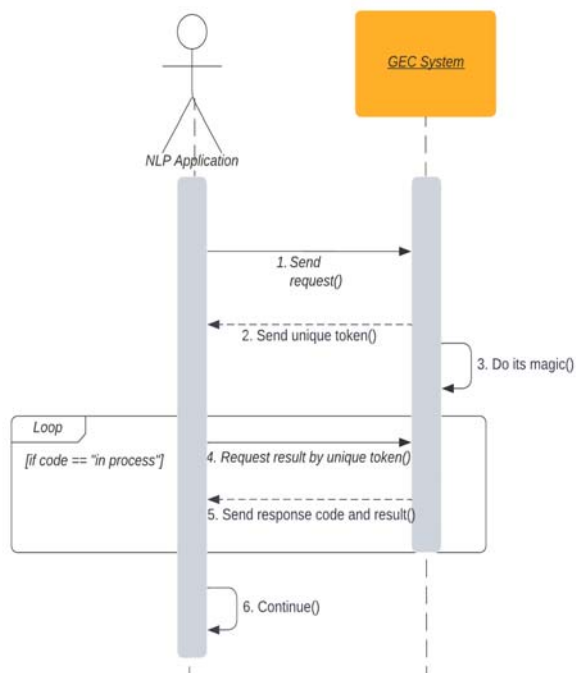


Рисунок 5 – Діаграма послідовності для другого режиму роботи

1. Надіслати запит (Send request).
2. Надіслати унікальний маркер (Send unique token).
3. Опрацювання запитів (Do its magic) з циклом (Loop) типу [якщо код == “в процесі”] ([if code == “in process”]).
4. Результат запиту за унікальним маркером (Request result by unique token)
5. Надіслати код відповіді та результат (Send response code and result).
6. Продовжити (Continue).

Об’єкт GEC System позначає сукупну систему граматичної корекції разом з усіма складовими, включаючи графічний інтерфейс, сервер для обробки даних, попередньо навчені ML-моделі. Повідомлення 3. Do its magic(), у свою чергу, позначає повний аналіз тексту та застосування необхідних виправлень.

GEC System надає кожному запиту унікальний токен, за допомогою якого можна отримати результати після завершення системою опрацювання запиту. Оскільки, в залежності від навантаження на систему, обробка даних може зайняти певний час, то NLP додаток має з певним інтервалом відправляти новий запит на отримання результатів. NLP додаток продовжує свою діяльність після отримання виправлених текстів. На діаграмі діяльності (рис. 6) зображений один із-поміж множини можливих підходів до розробки системи автоматичного виправлення граматичних помилок. Алгоритм:

1. Якщо доступний мовний словник [Language dictionary is available], то застосувати виправлення орфографії, що не запам’ятовується (Apply low-recall spelling correction).
2. Якщо мовний словник недоступний [Language dictionary is not available], то застосувати перевірку основних правил (Apply check of basic rules).
3. Якщо доступний синтаксичний аналізатор залежності/вибірчої групи [Dependency/Constituency parser is available], то застосувати перевірку більш конкретних правил (Apply check of more specific rules).
4. Якщо правила можуть охоплювати всі випадки [Rules can cover all cases], то використати підхід, заснований на правилах (Use a rule-based approach).
5. Якщо правила не можуть охоплювати всі випадки [rules can’t cover all cases] і/або якщо парсер недоступний [parser is not available], то перевірити або визначити розмір корпусу.
6. Якщо достатньо великі корпуси недоступні [large enough corpora is not available], то застосувати методи збільшення/генерування даних (Apply data augmentation / generation techniques).
7. доступні великі корпуси [large corpora is available], то визначити характеристики корпусу.
8. Якщо доступні великі анотовані корпуси [large annotated corpora is available], то використати підхід позначення послідовності (Use sequence tagging approach).
9. Якщо доступні великі паралельні корпуси [large parallel corpora is available], то використати підхід

нейронного машинного перекладу (Use neural machine translation approach).

10. Якщо доступні великі неанотовані корпуси [large unannotated corpora is available], то використати підхід на основі статистики (Use statistical-based approach).

Перш за все, якщо словник слів (або правил утворення слів) є доступним для певної мови, пропонується застосовувати його для виправлення орфографічних помилок. Варто зазначити, що модуль для автоматичного виправлення орфографічних помилок має мати низькі показники повноти (recall).

Значення повноти також можна інтерпретувати як відношення кількості правильно визначених позитивних випадків до усіх істинно позитивних випадків, тобто як частку загального числа позитивних зразків, що було знайдено.

Значення влучності (precision), у свою чергу, можна оцінити як частку правильно визначених істинно позитивних випадків.

В залежності від якості класифікатора та значення порогу прийняття рішення значення влучності буде залежним від значення повноти (рис. 7).

Низькі показники повноти (і високі – влучності) позначають те, що класифікатор (у даному випадку функція, що визначає орфографічні помилки) має бути достатньо «впевненим» у своєму «рішенні» про виправлення слова саме для того, щоб заміна не могла причинити зміну семантики речення.

Після перевірки і виправлення орфографії важливим є застосування перевірки базових правил граматики. Приклади для української мови: чергування «у-в», правопис «пів-», правопис апострофа та м'якого знака тощо.

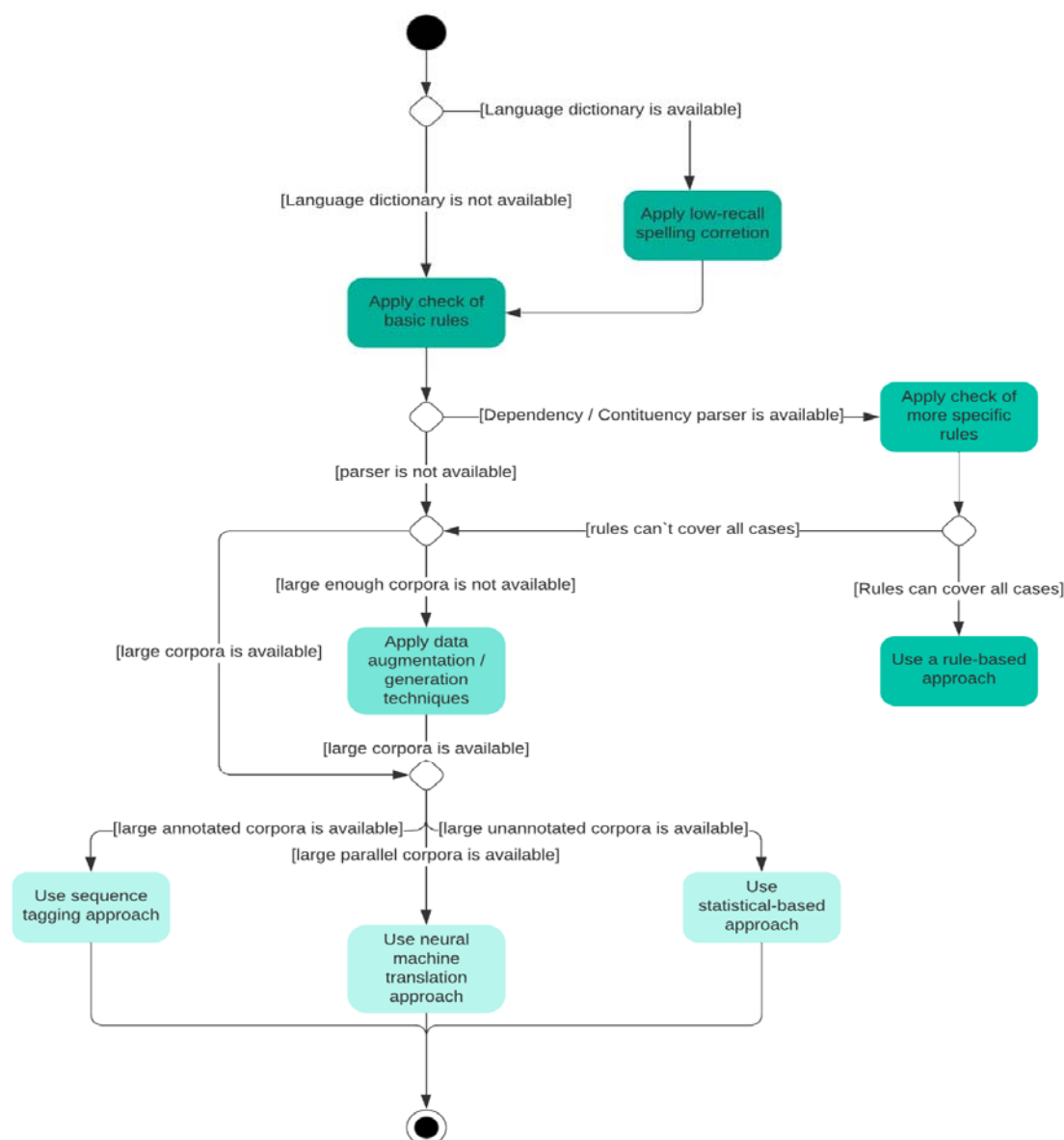


Рисунок 6 – Діаграма діяльності процесу розробки ГЕС-системи

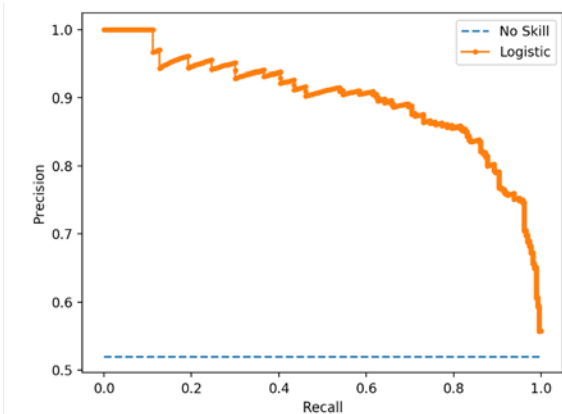


Рисунок 7 – Приклад кривої Precision-Recall

В залежності від того, чи є доступними парсери залежностей, груп слів, іменованих сутностей, частин мови тощо, можна застосовувати перевірку додаткових правил на основі отриманих результатів. Якщо теоретично можливий ітеративний розвиток системи і впроваджені правила можуть пояснити усі можливі помилки певної мови, використання методів і алгоритмів не є необхідним.

У випадку, якщо перевірки додаткових правил не є достатньо (особливо у випадку морфологічно багатих мов), для виправлення залишкових помилок можливим є застосування ML-методів. Якщо навчальних даних не є достатньо або вони взагалі не є доступними, для отримання паралельного / анотованого корпусу застосовують методи аугментації та генерації даних. Також, в залежності від того, чи є дані анотованими або корпус містить набір паралельних текстів, використовують відповідні підходи до розробки і навчання моделі. У випадку достатньо великих і якісних неанотованих корпусів також можна побудувати мовну модель на основі статистичних методів.

На діаграмі розгортання GEC-системи (рис. 8) зображено 3 основні процесори на основі трьох пристроїв/програм, зокрема : Пристрій користувача (:User Device), Сервер БД (:DB Server) та Сервер опрацювання даних (:Data Processing Server). :User Device застосовує програму/браузер (Program/Browser). :DB Server використовує База даних NoSQL (NoSQL Database). :Data Processing Server побудований на основі реалізації ML-моделі (ML model), словника (Dictionary), модуля перевірки правил (Rule-check module) та модуля огляду користувачів (User review module).

Користувач має доступ до функціоналу системи за допомогою графічного інтерфейсу завантаженої програми, розширення або онлайн-платформи. У БД зберігається інформація про користувачів, приклади запропонованих виправлень, документи. Така БД за своїм типом не є реляційною.

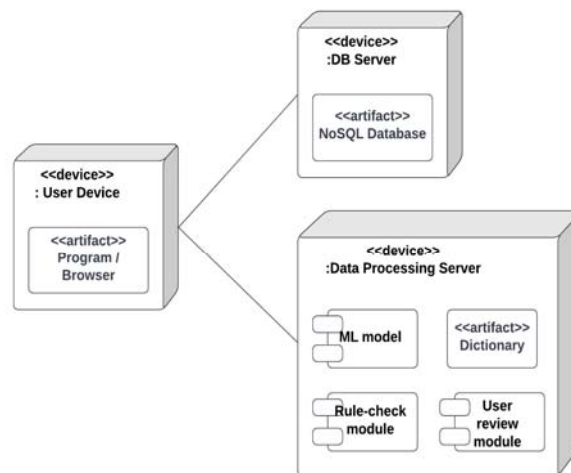


Рисунок 8 – Діаграма розгортання GEC-системи

Сервер обробки даних містить модулі ML-моделі, перевірки правил і перевірки запропонованих користувачами виправлень. Остання може застосовуватись для додаткового налаштування загальної системи перевірки та виправлення.

Альтернативним варіантом схеми розгортання є використання власних обчислювальних ресурсів персонального комп'ютера для аналізу даних, однак, з врахуванням того, що для отримання достовірного результату необхідно застосувати кілька етапів перевірки, особливо з використанням ML-методів, така архітектура не є доцільною.

Вибір набору даних. Набір україномовних UA-GEC анотованих текстів презентований дослідниками з Grammarly. Цей набір даних був професійно анотований для GEC та редагування україномовних текстів. Наскільки відомо на час публікації, це перший корпус GEC з української мови. Дослідники зібрали тексти з помилками (20 715 речень – 328 779 tokenів) від різноманітних авторів, у тому числі носіїв мови. Дані охоплюють найрізноманітніші сфери письма, від текстових чатів і есе до офіційного письма. Професійні коректори виправляли та анотували корпус на помилки, пов'язані з вільним мовленням, граматикою, пунктуацією та орфографією. Даний корпус можна використовувати для розробки та оцінки систем GEC українською мовою і дослідження морфологічно багатих мов.

Методи. Основні підходи виявлення помилок, на яких зосереджені академічні дослідження, можна розділити на такі категорії: методи, засновані на правилах; методи, засновані на синтаксичному аналізі речень; статистичне моделювання; класичні ML-методи; МТ на основі глибинного навчання. Для покращення якості системи використовують різні методи генерації синтетичних навчальних даних.

МТ. Поява у 2018 році нової архітектури глибинного навчання Transformer, що містить у собі механізм уваги, дозволила значно спростити розробку мовних моделей, що тепер може відбуватися без

застосування експертних знань домену, у даному випадку – лінгвістики. Більша частина досліджень застосовує модель Transformer для впровадження нейронного МТ тексту із помилками у його правильний варіант. Значним недоліком такого підходу для розробки моделей для різних мов є необхідність у великих корпусах анотованих або паралельних даних.

Transformer – модель глибинного навчання, особливістю якої є застосування механізму уваги, що роздільно зважує важливість кожної частини даних входу.

Будівельними блоками трансформера є вузли масштабованої скалярнодобуткової уваги (scaled dot-product attention units). Коли моделі Transformer передають речення, ваги уваги між усіма лексемами обчислюються одночасно. Вузол уваги виробляє вкладення для кожної лексеми в контексті, які містять інформацію про саму лексему, разом зі зваженим поєднанням інших релевантних лексем, кожен з яких зважено за вагою уваги до неї.

Альтернативою архітектурі Transformer є згорткові або рекурентні нейронні мережі архітектури encoder-decoder.

– Згорткові нейронні мережі (CNN, Convolutional Neural Network)

CNN складається з шарів входу та виходу, а також із декількох прихованих шарів. Приховані шари CNN зазвичай складаються зі згорткових шарів, агрегувальних шарів, повноз'єднаних шарів та шарів нормалізації. Згорткові шари застосовують до вхідних даних операцію згортки, передаючи результат до наступного шару.

Згорткові мережі можуть включати шари локального або глобального агрегування, які об'єднують виходи кластерів нейронів одного шару до наступного шару. Наприклад, максимізаційне агрегування використовує максимальне значення з кожного з кластерів нейронів попереднього шару. Іншим прикладом є усереднене агрегування, що використовує усереднене значення з кожного кластеру нейронів попереднього шару. Повноз'єднані шари з'єднують кожен нейрон одного шару з кожним нейроном наступного шару.

– Рекурентні нейронні мережі (Recurrent Neural Networks, RNN)

Рекурентні нейронні мережі це мережі, що містять зворотні зв'язки і дозволяють зберігати інформацію. Наявність зворотного зв'язку дозволяє передавати інформацію від одного кроку навчання мережі до іншого.

Одним з різновидів RNN є LSTM-мережі. LSTM (Long Short Term Memory) – це RNN, здатні до навчання довготривалих залежностей. LSTM-мережі складаються з повторюваних елементів. Кожен такий елемент містить чотири шари і відрізняється тим, що має мірку довгої короткочасної пам'яті.

Можливість LSTM-мереж успішно вивчати дані з довготривалими залежностями робить їх

оптимальним вибором для розв'язання задач, у котрих як вхідна, так і вихідна інформація представляються у вигляді послідовностей деяких елементів (наприклад, літер, слів, речень).

Sequence tagging. Окрім методу нейронного перекладу, ML-методи у задачі виявлення та виправлення орфографічних помилок також застосовують для передбачення міток, що позначають дію, яку необхідно виконати над кожним токеном. В залежності від типу помилок, кількості наявних класів міток та загальної складності завдання використовують як глибинні ML-методи, так і класичного.

До класичних ML-методів із вчителем відносяться такі алгоритми: логістична регресія, дерево рішень, найвний байєсів класифікатор, метод опорних векторів, k-найближчих сусідів, випадковий ліс тощо.

– Логістична регресія

Логістична регресія використовується для завдань бінарної класифікації, тобто коли на виході потрібно отримати відповідь, до якого з двох класів належить об'єкт. Логістична функція перетворює будь-яке значення в число в межах від 0 до 1. Так і передбачається ймовірність належності до одного чи другого класу. Тобто, на відміну від логістичної регресії, цей метод не передбачає значення числової змінної, виходячи з вибірки вихідних значень, а замість цього значенням функції є ймовірність того, що це початкове значення належить до певного класу.

– Дерево рішень

Дерево рішень буде моделі класифікації або регресії у вигляді деревоподібної структури. Метод розбиває набір даних на менші й менші підмножини. Кінцевий результат – дерево з вузлами рішення і кінцевими вузлами. Що глибше дерево, то складніші правила ухвалення рішень і точніша модель.

Однією з основних переваг дерев рішень є їхня інтуїтивність: класифікаційна модель, що представлена у вигляді дерева рішень, легко інтерпретується користувачем та спрощує розуміння завдання, яке потрібно вирішити, тому що дозволяє зрозуміти, чому конкретний об'єкт належить до відповідного класу. Ще однією важливою перевагою дерев рішень є їхня універсальність для вирішення задач класифікації та регресії. Але є в алгоритму і недоліки. По-перше, це нестабільність процесу: невеликі зміни в наборі даних можуть призводити до побудови іншого дерева. Це пов'язано з ієрархічністю дерева: зміни в вузлах на верхньому рівні можуть призвести до змін у всьому дереві нижче. По-друге, це складність контролю розміру дерева, що є критичним фактором, який зумовлює якість вирішення завдання.

– Наївний Байєсів класифікатор

Наївний метод Байєса – це набір методів класифікації, заснованих на теоремі Байєса. Модель складається з двох типів ймовірностей, які розраховуються за допомогою тренувальних даних: ймовірність кожного класу і умовна ймовірність для

кожного класу при кожному значенні  $x$ . Наївний байесів класифікатор називається наївним, тому що алгоритм передбачає, що кожна вхідна змінна є незалежною. Це припущення не завжди відповідає реальним даним. Проте даний алгоритм вельми ефективний для цілого ряду складних завдань на зразок класифікації спаму або розпізнавання рукописних цифр.

#### – Метод опорних векторів

Метод опорних векторів використовується для задач класифікації тексту, як-от призначення категорії і виявлення спаму. Метод опорних векторів – один із найбільш популярних для класичної класифікації, тому що він простий та швидкий. Виходячи з того, що об'єкт, який перебуває в  $N$ -вимірному просторі, належить до одного з двох класів, метод опорних векторів буде гіперплощиною з розмірністю  $(N - 1)$ , щоб всі об'єкти виявилися в одній з двох груп. Що далі від гіперплощини лежать точки даних, то більше впевненість в тому, що вони були правильно класифіковані. Що менше схожих ознак між даними, то менша ймовірність належності до одного класу.

#### – Метод $k$ -найближчих сусідів

Для алгоритму  $k$ -найближчих сусідів передбачення для нової точки робиться шляхом пошуку  $k$  найближчих сусідів в наборі даних і підсумовування вихідної змінної для цих  $k$  примірників. Ідея найближчих сусідів може погано працювати з багатовимірними даними, що негативно позначиться на ефективності алгоритму при вирішенні задачі.

#### – Випадковий ліс

Випадковий ліс належить до сімейства ансамблевих алгоритмів. Як зрозуміло з його назви, він містить велику кількість окремих дерев рішень, які діють як ансамбль. Кожне окреме дерево у випадковому лісі обчислює прогноз класифікації, і клас із найбільшою кількістю голосів стає прогнозом моделі. Результатом обирається той варіант, який випадав найчастіше. Велика кількість некорельованих дерев рішень, тобто тих, що знаходять рішення незалежно одне від одного й діють спільно, перевершить будь-яке рішення, отримане одним деревом рішення.

Саме некорельовані моделі можуть створювати ансамблеві прогнози, які є точнішими, ніж будь-який з окремо взятих прогнозів. Причиною такого ефекту є те, що кожне дерево рішень «захищає» одне одного від індивідуальних помилок: хоча деякі дерева рішень можуть бути неправильними, більшість із них будуть правильними, тому як група дерев буде рухатися в правильному напрямку.

До переваг методу можна віднести можливість ефективно обробляти дані з великою кількістю ознак і класів та високу масштабованість. Серед недоліків можна виділити великий розмір моделей, що будуються. Що більша модель, то вища її обчислювальна складність та швидкість знаходження рішень.

#### – AdaBoost

Бустінг – це сімейство ансамблевих алгоритмів, суть яких полягає в створенні сильного класифікатора на основі декількох слабких. Для цього спочатку створюється одна модель, потім інша модель, яка намагається виправити помилки в першій. Моделі додаються до тих пір, поки тренувальні дані не будуть ідеально класифіковані або поки не буде перевищено максимальну кількість моделей.

AdaBoost використовують разом з короткими деревами рішень. Після створення першого дерева перевіряється його ефективність на кожному тренувальному об'єкті, щоб зрозуміти, скільки уваги має приділити наступне дерево всіх об'єктах. Тим даним, які складно передбачити, дається більшу вагу, а тим, які легко передбачити, – меншу. Моделі створюються послідовно одна за одною, і кожна з них оновлює ваги для наступного дерева. Після побудови всіх дерев робляться прогнози для нових даних, і ефективність кожного дерева залежить від того, наскільки точним воно було на тренувальних даних.

Для обрання найкращого методу класифікації для його подальшого застосування необхідно порівняти показники точності, повноти, влучності,  $F$ -міри для всіх вищезазначених алгоритмів.

Морфологічні парсери. Перевірка правил відмінювання та узгодження слів у реченні може бути заснована на парсингу дерев залежності (Dependency Parsing) або приналежності (Constituency Parsing).

Термін парсинг залежностей відноситься до процесу дослідження залежностей між фразами речення з метою визначення його граматичної структури. Речення ділиться на багато розділів здебільшого на основі цього. Процес заснований на припущенні, що між кожною мовною одиницею в реченні існує прямий зв'язок. Ці гіперпосилання називаються залежностями. На відміну від аналізу залежностей, парсинг приналежностей розподіляє речення на фрази певного типу (іменникова, дієслівна групи тощо). У цьому випадку термінальний вузол – це мовна одиниця або фраза, яка має батьківський вузол та тег частини мови.

Парсинг залежностей та приналежностей слів у реченнях певних мов можна виконати за допомогою бібліотеки для NLP-NLTK, однак вона не підтримує аналіз україномовних текстів.

Підтримка української мови. Бібліотека мови Python `rumorphy2` підтримує обробку української мови. `rumorphy2` аналізує частину мови, число, відмінок, час, лему та стем заданого слова. Морфологічний парсер заснований на словниках `OpenCorpora`, що конвертовані до формату XML. Користувачі також мають змогу додавати власні слова та правила, це дає можливість проводити морфологічний аналіз текстів певної предметної області без зміни вихідного коду `rumorphy2` та адаптувати `rumorphy2` для роботи з іншими мовами.

Мова і бібліотеки. Для ML найпопулярнішими мовами є Python, R, Java, C++. Перевагою Python з-



поміж інших мов саме для створення системи визначення емоційного забарвлення текстового контенту є його підтримка великої кількості бібліотек:

- для роботи з класичними ML-методами: Scikit-Learn;
- для створення штучних нейронних мереж, глибинного навчання: TensorFlow, Keras, PyTorch;
- для NLP: NLTK, spaCy, WordNet;
- для роботи із масивами, матрицями: NumPy;
- роботи з таблицями: pandas;
- візуалізації даних (в тому числі, й інтерактивної): Matplotlib, seaborn, Plotly.

Бібліотека Scikit-Learn підтримує TSP, зменшення розмірності даних, вибір ML-моделей для регресії, класифікації або кластерного аналізу. Однак Scikit-Learn не має комплексної підтримки для створення моделей глибинного навчання.

Для створення штучних нейронних мереж була обрана бібліотека Keras, що слугує високорівневим API для TensorFlow 2. Keras дозволяє будувати послідовні моделі у вигляді графу, вершинами якого є шари (layers) певного типу із заданою кількістю вузлів. Також, за допомогою Keras можна поєднувати результати роботи кількох окремих частин нейронної мережі для їх подальшої обробки, така структура не є лінійною.

TensorFlow дає можливість імпортувати навчену ML-модель для її подальшого використання у інших програмах. TensorFlow також підтримує виконання низькорівневих операцій із тензорами за допомогою центральних процесорів, графічних процесорів та тензорних блоків обробки.

Бібліотека NLTK у цьому дослідженні застосовується для попередньої обробки тексту: токенизації, видалення стоп-слів, стемінгу, лематизації. Також за допомогою функцій з цієї бібліотеки можна виявляти найпопулярніші n-грами і частини мови окремих токенів, розпізнавати іменовані сутності тощо.

До додаткових бібліотек, що спрощують роботу із природньою мовою, належать Regex та emoji – для використання регулярних виразів і заміни емоджі словами відповідно.

Середовище розробки. Для роботи над задачами з дослідження даних і застосування ML зручним інструментом є Jupyter Notebook, який дозволяє запускати написаний код невеликими фрагментами – комірками. Одним із онлайн-сервісів, що надає змогу використовувати Jupyter Notebook без локального встановлення, є Google Colab. Цей сервіс дає можливість використовувати графічні процесори GPU і TPU, що значно пришвидшують навчання штучних нейронних мереж.

#### 4 ЕКСПЕРИМЕНТИ

Опрацювання набору даних. Обраний набір даних UA-GEC [15] містить 850 і 160 текстів у навчальній і тестовій вибірці відповідно. За даними авторів, набір

даних містить усього 20 715 речень (як з помилками, так і без). На момент написання роботи вбудовані методи ітерації по корпусу UA-GEC дозволяють лише отримувати повні документи, а не окремі речення.

Текст анований за наступним форматом:

```
{like=>likes:::error_type=Grammar} turtles.
```

Для опрацювання та розділення текстів була застосована бібліотека ReGex. Регулярні вирази для розподілу текстів на речення та виявлення помилок мають наступний формат:

```
split_pattern = r'\n+'  
additional_split_pattern = r'(?<=[^А-Я].[?!(\.\.\.])  
+(?=[А-Я""'])'  
error_pattern =  
r'\{([^\{]*\{.*\})=>([^\{]*\{.*\}):::error_type=([^\{]*\})\}'
```

Правильні пари не видаляються, однак для подальшого опрацювання були обрані речення, де кількість токенів є більшою за 4 (два з них – спеціальні токени початку і кінця речення).

Усього в результаті розподілу текстів за вищезазначеними виразами отримуємо 15 599 і 2205 речень у навчальній та тестовій вибірках відповідно.

Попередньо навчені нейронні мережі.

Морфологічна складність української мови зумовлює потребу у поєднанні різних GEC-методів.

Один із модулів запропонованої системи – опрацювання речення за допомогою штучних нейронних мереж. Навчання нейронної мережі «з нуля» з випадково ініціалізованими вагами потребує паралельних корпусів великих обсягів. Єдиний на момент написання роботи україномовний набір даних містить усього тисячу текстів, що є критично мало для навчання мовної моделі для обробки морфологічно багатих мов.

Тому, навчання нейронної мережі GEC-завданню вимагає її попереднього налаштування за допомогою великих корпусів. У цьому випадку існує кілька попередньо налаштованих моделей, що підтримують українську і параметри яких є доступними у відкритому доступі: RoBERTa від YouScan [48], MT-моделі (від Google – MT5 [49], від Facebook – M2M100 [50] та mBART-50 [51]).

Модель архітектури енкодер-декодер на основі RoBERTa. У 2020 році команда дослідників із YouScan презентувала нейронну мережу «Ukrainian Roberta» [48], що була попередньо натренована на україномовних текстах загальним обсягом 2,5 мільярди слів. «Ukrainian Roberta» має 125 мільйонів параметрів, що налаштовуються, її навчання тривало 85 годин.

Складністю використання цієї нейронної мережі є те, що вона за своєю архітектурою є тільки енкодером і використовується для отримання контекстних векторних вкладень слів. «Ukrainian Roberta» була навчена вирішувати завдання передбачення замаскованого токена у реченні (рис. 9). Попереднє навчання нейронної мережі дозволяє їй отримати попереднє «знання» граматики мови.

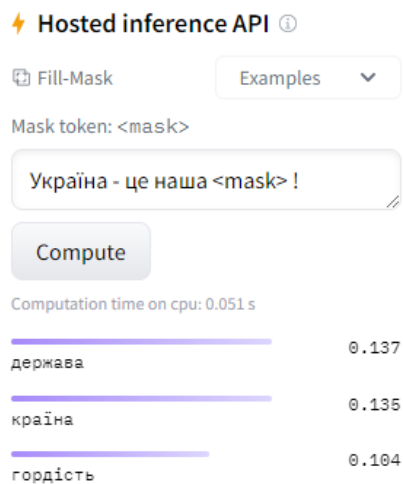


Рисунок 9 – Приклад передбачення замаскованого токена

BERT – модель глибинного навчання для NLP архітектури Transformer. Базова модель BERT має 110 мільйонів, розширена версія – 345. Особливістю архітектури Transformer є наявність механізму уваги, завдяки чому дані можуть бути оброблені одночасно (на противагу рекурентним нейронним мережам, де дані сприймаються послідовно). Окрім того, особливістю BERT є попереднє навчання нейронної мережі для вирішення двох завдань: передбачення певного слова у реченні і визначення того, чи є друге речення логічним продовженням першого. Попереднє навчання та механізм уваги дають змогу отримати контекстні векторні вкладення слів.

Попередньо навчену нейронну мережу RoBERTa можна вважати покращеним аналогом BERT: її архітектура є аналогічною до архітектури BERT, основна відмінність полягає у підборі гіперпараметрів під час попереднього навчання, збільшенні об'єму навчального корпусу, застосуванні динамічного маскування токенів для задачі передбачення слова у реченні: слово, яке необхідно передбачити у певному реченні, змінюється з кожною епохою. Окрім того, у цьому випадку відсутнє передбачення того, чи є друге речення логічним продовженням першого. Наприклад, для класифікації речень додається шар повнозв'язних нейронів, де кількість нейронів відповідає наявним класам [52].

Нейронні мережі типу денкодер утворюють контекстні векторні представлення слів у реченні фіксованої довжини незалежно від довжини вхідного повідомлення. У свою чергу, BERT є складовою нейронних мереж типу енкодер-декодер, що використовуються для завдань sequence-to-sequence. Декодер, у свою чергу, генерує токени послідовно, кожний наступний токен залежить від попередніх. Застосування попередньо навченої «Ukrainian Roberta» є можливим за аналогією до підходу, запропонованого у [53]. У випадку ініціалізації BERT у якості декодера, для передбачення наступного слова на основі контекстного векторного вкладення вхідних токенів,

необхідно додати шар перехресної уваги (cross-attention layer) із випадковим чином згенерованими вагами. Окрім того, для передбачення розподілу ймовірностей наступного слова використовується LM Head (шар або шари повнозв'язних нейронів, де кількість вихідних нейронів відповідає кількості токенів у словнику). Вагові коефіцієнти LM Head ініціалізуються ваговими коефіцієнтами шару векторного вкладення BERT  $W_{emb}$ . Також, двонаправлений механізм уваги BERT необхідно замінити на однонаправлений для генерації токена в залежності лише від попередніх [52, 57]. Словник токенів, що відповідає попередньо навченій моделі «Ukrainian Roberta», містить 52 тис. елементів, де перші записи є спеціальними токенами, розділовими знаками та найбільш поширеними словами (рис. 10). Оскільки словники токенів є унікальними для кожної попередньо навченої моделі, TRP теж має свої особливості. Після токенизації текстів і речень у навчальній вибірці розподіл кількості токенів має наступний вигляд, поданий на рис. 11. Максимальна довжина речення – 100 токенів, довші речення обрізаються.

Нейронна мережа навчалась протягом 15 епох з наступними гіперпараметрами:

- `n_epochs = 15`
- `batch_size = 8`
- `lr = 0.00001`

Після 15-ї епохи значення функції втрат на тестовій вибірці збільшується (рис. 12), тому 15-ть епох є найбільш оптимальним значенням у даному випадку.

Налаштування попередньо навченої MT-моделі. Facebook та Google надали відкритий доступ до ваг попередньо навчених нейронних мереж для MT.



Рисунок 10 – Словник токенів «Ukrainian Roberta»

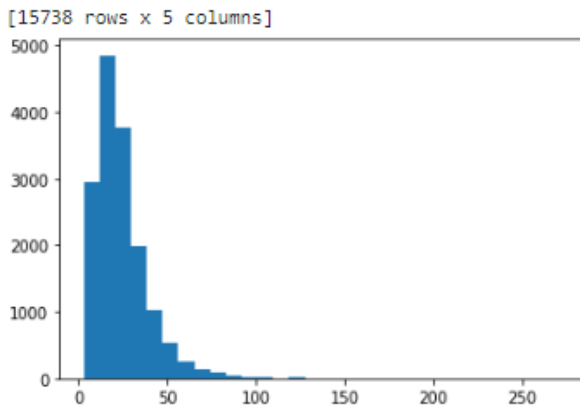


Рисунок 11 – Розподіл кількості токенів у навчальній вибірці

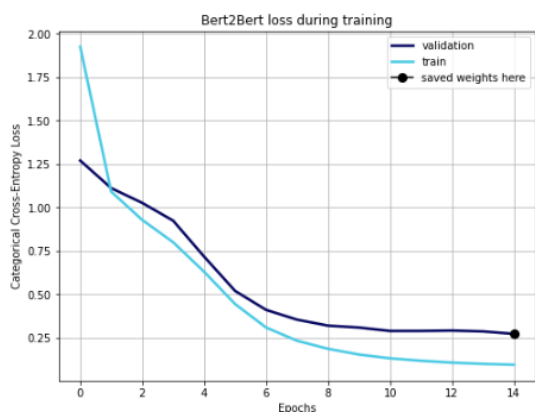


Рисунок 12 – Крива функції втрат при навчанні NN

mT5 [49] заснована на архітектурі Transformer і була попередньо навчена MT-завданням із застосуванням корпусу, заснованого на даних Common Crawl. Усього нейронна мережа була попередньо натренована на текстах, написаних 101 мовою, серед яких є й українська. Базова версія, що застосовується у цьому дослідженні, має 580 мільйонів параметрів, що налаштовується, версія mT5-XXL – 13 мільярдів параметрів.

У зв'язку з обмеженням наявних обчислювальних ресурсів, найбільша довжина речення складає 20 токенів, batch size – 2 записи. Варто зауважити, що, оскільки нейронна мережа була попередньо натренована за допомогою текстів, написаних різними мовами, то відповідний словник токенів радше містить частини україномовних слів, аніж повні слова. Тому одне слово кодується більшою кількістю токенів, а у контексті обмеженої довжини речення це може призвести до некоректного представлення навчальних прикладів і їх значення, що не є повністю завершеним. Нейронна мережа навчалась протягом 10 епох, найнижче значення функції втрат на тестовій вибірці було досягнуто на 4 епосі, тоді ж і були збережені ваги нейронної мережі.

Відповідно до [50], M2M100 має низку переваг у порівнянні із попередніми MT-моделями. За ствердженнями дослідників, попередні MT-моделі © Холодна Н. М., Висоцька В. А., 2023  
DOI 10.15588/1607-3274-2023-1-12

навчаються на «англо-центричних» даних, тобто для перекладу між двома мовами такі моделі навчаються перекладу з першої мови на англійську, потім – з англійської на другу мову. Такий підхід може призвести до втрати початкового значення речення, і M2M100 – перша нейронна мережа для MT, що була навчена прямому перекладу з першої мови на другу без проміжного перекладу на англійську. Такий підхід дозволив отримати найкращі результати метрики BLEU у порівнянні з іншими дослідженнями, збільшивши показник на 10 пунктів. Усього нейронна підтримує 2 200 напрямків перекладу – у 10 разів більше, ніж попередня англо-центрична мовна модель. NN попередньо натренована на тестах, написаних 100 мовами, у тому числі й українською. Базова версія, що застосовується у дослідженні, містить 418 мільйонів параметрів, найбільша – 15 мільярдів. Після розрахунку кількості токенів для текстів навчальної вибірки отримуємо наступний розподіл, поданий на рис. 13.

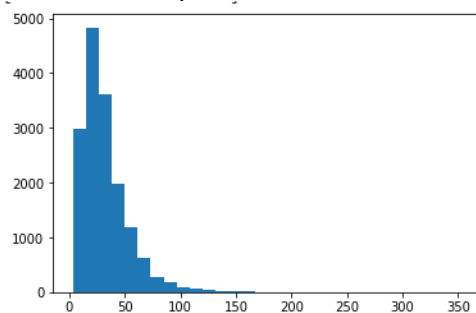


Рисунок 13 – Розподіл кількості токенів у навчальній вибірці

Аналогічно до попереднього випадку, максимальна кількість токенів є обмеженою у зв'язку з обмеженими обчислювальними ресурсами. В цьому випадку, також можливе некоректне представлення навчальних прикладів. Нейронна мережа навчалась протягом 10 епох, однак найменше значення функції втрат на тестовій вибірці було досягнуто на другій епосі (рис. 14). Для порівняння також було проведено навчання нейронної мережі протягом 10 епох, ваги були збережені після останньої.

Варто зауважити, що на якість генерації тексту впливає низка параметрів, а саме:

- Значення гіперпараметрів.
- Обсяг навчальної вибірки.
- Кількість навчальних епох.
- Архітектура і кількість параметрів нейронної мережі.
- Кількість токенів вхідного тексту.
- Навчальна вибірка (чи застосовувались валідаційні дані або був використаний весь корпус).

Приклад навчання NN MT5. Імпорт бібліотек. Основними бібліотеками у даному проекті є PyTorch, HuggingFace Transformers, Regex, pandas (рис. 15).

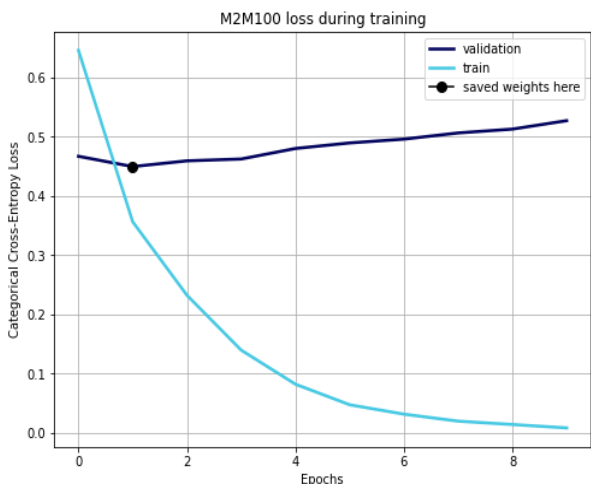


Рисунок 14 – Крива функції втрат при навчанні NN

```
##### IMPORTS #####
from collections import defaultdict
from ua_gec import Corpus
from ua_gec.corpus import Document

import pandas as pd
import regex as re
import copy
import numpy as np
import matplotlib.pyplot as plt

from transformers import AdamW, AutoModelForSeq2SeqLM, AutoTokenizer
from transformers import get_linear_schedule_with_warmup
from transformers import Seq2SeqTrainer, Seq2SeqTrainingArguments

from IPython.core.display import display, HTML
from tqdm import tqdm_notebook

import torch
from torch.utils.data import Dataset, DataLoader
```

Рисунок 15 – Імпорт бібліотек

Константи на налаштування. Встановлюємо такі налаштування: регулярні виразу для розділення текстів на речення, регулярний вираз для пошуку анотації помилок, гіперпараметри (рис. 16). Завантаження попередньо навчено моделі. Завантаження попередньо навченої NN для її подальшого налаштування відбувається за допомогою кількох команд (рис. 17). Створення PyTorch Dataset. Перед створенням класу Датасет та його об'єктів необхідно провести попереднє дослідження даних, після цього агрегуючи ці етапи (рис. 18). Додаткові налаштування подані на рис. 19. При використанні PyTorch необхідно кожного разу реалізувати власну функцію навчання NN (рис. 20).

```
##### SETTINGS #####
split_pattern = r'\n+'
error_pattern = r'\{([\{\}\*\(\)\*\})::error_type=([\{\}\*\})\}'
additional_split_pattern = r'(?<=[^A-R].[!?\(\.\.\.\.])+(?=[A-R]*)'

regex_special_symbols = {
    '(': '\(',
    ')': '\)',
    '?': '\?',
    '[': '\[',
    ']': '\]',
    '*': '\*',
    '$': '\$'
}

model_name = 'google/mt5-base'

n_epochs = 25
batch_size = 4
lr = 0.00001
max_token_length = 50

device = torch.device("cuda" if torch.cuda.is_available() else "cpu")
```

Рисунок 16 – Змінні із налаштуваннями

```
##### LOAD PRETRAINED MODEL #####
tokenizer = AutoTokenizer.from_pretrained(model_name)
model = AutoModelForSeq2SeqLM.from_pretrained(model_name)
model = model.to(device)
```

Рисунок 17 – Завантаження попередньо навченої NN

```
train_dataset = Parallel_Dataset('train', tokenizer, max_token_length)
test_dataset = Parallel_Dataset('test', tokenizer, max_token_length)
```

Рисунок 18 – Завантаження датасету

```
##### ADDITIONAL #####
test_loader = DataLoader(test_dataset, batch_size=batch_size)
train_loader = DataLoader(train_dataset, batch_size=batch_size)

n_batches = int(np.ceil(len(train_dataset)) / batch_size)
n_batches_test = int(np.ceil(len(test_dataset)) / batch_size)

total_steps = n_epochs * n_batches

n_warmup_steps = int(total_steps * 0.01)

optimizer = torch.optim.AdamW(model.parameters(), lr=lr)
scheduler = get_linear_schedule_with_warmup(optimizer,
                                             n_warmup_steps,
                                             total_steps)
```

Рисунок 19 – Додаткові налаштування

```
model, history = train(model, optimizer, scheduler,
                       n_epochs, len(train_dataset), len(test_dataset),
                       train_loader, test_loader)
```

Рисунок 20 – Виклик функції навчання NN

Збереження ваг NN: `torch.save(model, 'mt5.pt')`.  
 Отримання передбачення подано на рис. 21.

```
def predict(sentence):
    inputs = tokenizer(sentence, padding="max_length", truncation=True, max_length=200, return_tensors="pt")
    input_ids = inputs.input_ids.to("cuda")
    attention_mask = inputs.attention_mask.to("cuda")

    outputs = model.generate(input_ids, attention_mask=attention_mask)
    print(outputs)

    output_str = tokenizer.batch_decode(outputs, skip_special_tokens=True)

    displayHTML('<h3>Input sentence</h3>')
    displayHTML(f'<p>{sentence}</p>')
    displayHTML('<h3>Result</h3>')
    displayHTML(f'<p>{output_str[0]}</p>')

predict('учора був чудовий день')
```

Рисунок 21 – Функція отримання передбачення

## 5 РЕЗУЛЬТАТИ

Для демонстрації роботи нейронних мереж використані власні приклади та завдання Зовнішнього Незалежного Оцінювання. Отримані результати: «Ukrainian Roberta» encoder-decoder. Нейронна мережа непогано розставляє розділові знаки у простому реченні (рис. 22).

```
predict('я й не думав що лінгвістика це легко')

tensor([[ 0, 0, 848, 462, 355, 18554, 16, 402, 17134, 14428,
         341, 622, 537, 4222, 4222, 2]])
['Я й не думав, що лінгвістика – це легко!']

predict('Лиш бід, нехай щезає від тебе! - радіє спокійно хлопець [...] - Нема з ким говорити!')

Input sentence
Лиш бід, нехай щезає від тебе! - радіє спокійно хлопець [...] - Нема з ким говорити!

Result
Лиш бід, хай щезає від тебе! — радіє спокійно хлопець. — Нема з ким говорити!
```

Рисунок 22 – Розставлення розділових знаків

Наступний приклад: NN правильно виправила розділові знаки, однак самостійно змінила речення, використовуючи відомі їй словосполучення (рис. 23). Інколи NN «не розуміє» суті поставленого завдання (рис. 24 – залишити правильне словосполучення у початковому вигляді). Інколи NN взагалі «не розуміє» суті поставленого завдання (рис. 25). Однак у цьому випадку речення має логічну структуру. На рис. 26 NN правильно змінила відмінок іменника з називного на кличний, однак гуаш – фарба, іменник жіночого роду, тут він не є правильно узгодженим, ймовірно тому, що аналогічних прикладів немає у навчальному наборі даних.

```
predict('Я цієї пісні раніше не чув, - сказав студентіві Василь: - Ви її всю знаєте?')

Input sentence
Я цієї пісні раніше не чув, - сказав студентіві Василь: - Ви її всю знаєте?

Result
Я цієї пісні раніше не чув, - сказав студент Василь Васильєв. - Ви її всю історію знаєте?

predict('«Це ж хто сказав, що одна ластівка не робить весни?» - переможно вигукнув Таран.')

Input sentence
«Це ж хто сказав, що одна ластівка не робить весни?» - переможно вигукнув Таран.

Result
«Це ж хто сказав, що одна ластівка не робить весни?» — вигукнув Таран.
```

Рисунок 23 – Зміна речення нейронною мережею

```
predict('кришталеве джерело')

Input sentence
кришталеве джерело

Result
Решталеве джерело — джерело джерело для того призначення
```

Рисунок 24 – Нерозуміння задачі NN

```
predict('експонат в музеї')

Input sentence
експонат в музеї

Result
Експонат в музеї "Фантом", де ви побували на музеї, де й музеї "спався на музеї".

predict('далека путь')

Input sentence
далека путь

Result
далека путь — це дорігій шлях
```

Рисунок 25 – Некоректна зміна речення

```
predict('Слухай-но, Остап, чи не міг би ти купити мені червоний гуаш?')

Input sentence
Слухай-но, Остап, чи не міг би ти купити мені червоний гуаш?

Result
Слухай-но, Остап, чи не міг би ти купити мені червоний гуаш?
```

Рисунок 25 – Зміна відмінку іменника

На рис. 26 NN правильно виправила числівник, але завершила речення своєю «творчістю». У прикладах, що містять лише слова або словосполучення, нейронна мережа «намагається» завершити думку, самостійно добудовуючи речення. Такий результат є прямим наслідком того, що приклади у навчальному наборі містять лише речення, а не словосполучення.

```
predict('п'ятнадцять')

Input sentence
п'ятнадцять

Result
П'ятнадцять років — це п'ятнадцять.
```

Рисунок 26 – «Творчість» NN

Однак, для перевірки слів і словосполучень в загальному застосовуються інші методи: перевірка за словником та перевірка простих правил (апостроф, чергування у-в, правопис пів-, напів- тощо).

Попередні результати тестування NN показують, що NN неправильно визначає суть поставленого завдання (виправлення помилок при збереженні початкового змісту речення / словосполучення), що може бути результатом відсутності окремих слів та словосполучень у навчальному наборі даних. На протигагу цьому, NN виправляє речення паттернами з набору навчальних даних, що навпаки може свідчити про її перенавчання. Окрім того, причиною таких помилок може бути архітектура моделі – RoBERTa оригінально застосовується як енкодер для отримання контекстних представлень токенів. На рис. 27 NN виправила «чим...тим» на «що...то». Деякі

мовознавці вважають, що цей сполучник не є власне українським, а суржиком. Прості речення є зазвичай простим завданням для NN (рис. 28), але є складності (рис. 29).

**Input sentence**

Один мудрець вдало підмітив: Чим більше пізнаєш людей, тим більше починаєш любити тварин.

**Result**

Один мудрець вдало підмітив: що більше пізнаєш людей, то більше починаєш любити тварин.

Рисунок 27 – Виправлення NN суржика

**Input sentence**  
Концерт тривав майже півтора години.

**Result**  
1 - Концерт тривав майже півтори години.  
2 - Конгрес тривав майже півтори години.  
3 - Концерт тривав майже півтора години.

**Input sentence**  
Хочу подякувати учасників.

**Result**  
1 - Хочу подякувати учасникам.  
2 - Конгрес тривав майже півтори години.  
3 - Концерт тривав майже півтора години.

**Input sentence**  
Водій автобуса сказав пасажиром щоб вони оплати проїзд.

**Result**  
1 - Водій автобуса сказав пасажиром, щоб вони оплати проїзд.  
2 - Водій автобуса сказав пасажиром, щоб вони оплати проїзду.  
3 - Водій автобуса сказав пасажиром, щоб вони оплатили проїзд.

**Input sentence**  
вона хотіла би навчатися малювання.

**Result**  
1 - Вона хотіла би навчатися малювати.  
2 - Він турботливий за відношення до майна.  
3 - Він турботливий до майна.

Рисунок 28 – Виправлення простих речень

**Input sentence**  
учора був чудове день народження.

**Result**  
Учора був чудовий день народження.

Рисунок 29 – Некоректне виправлення

Правильне речення є серед пропозицій моделі, однак проблемним є той факт, що більша ймовірність була призначена неправильному реченню (рис. 30). На рис. 31а NN правильно замінила словосполучення «за професії», однак слово «архітектор» замінене на найближче за значенням слово. Аналогічно, більшу ймовірність має неправильне речення (рис. 31б), але правильне виправлення є серед запропонованих.

У деяких реченнях mT5 пропонує більш правильні виправлення (рис. 32). На даних прикладах найкращу точність виправлення має MT-модель mT5, однак для остаточної оцінки і порівняння NN необхідний достатньо великий тестовий корпус даних.

**Input sentence**  
учора був чудове день народження.

**Result**  
1 - Учора був чудове день народження.  
2 - Учора був чудовий день народження.  
3 - учора був чудове день народження.

Рисунок 30 – Варіанти виправлення

**Input sentence**  
він за професії архітектор.

**Result**  
1 - Він за фахом інженер.  
2 - Він за професії інженер.  
3 - Він за спеціальністю інженер.

а б  
Рисунок 31 – Коректне виправлення Roberta mT5

<p><b>Input sentence</b> він за професії архітектор.</p> <p><b>Result</b> 1 - Він за фахом інженер. 2 - Він за професії інженер. 3 - Він за спеціальністю інженер.</p>	<p><b>Input sentence</b> він за професії архітектор.</p> <p><b>Result</b> 1 - Він за професії архітектор. 2 - Він за професією архітектор. 3 - Він – за професії архітектор.</p>
<p><b>Input sentence</b> Хочу подякувати учасників.</p> <p><b>Result</b> Хочу подякувати учасникам.</p>	<p><b>Input sentence</b> Хочу подякувати учасників.</p> <p><b>Result</b> 1 - Хочу подякувати учасникам. 2 - Хочу подякувати учасникам. 3 - Хочу подякувати учасників – Хочу подякувати учасників</p>
<p><b>Input sentence</b> це сама грша ситуація</p> <p><b>Result</b> 1 - Це сама грша ситуація 2 - Я сама грша ситуація 3 - — Це сама грша ситуація</p>	<p><b>Input sentence</b> це сама грша ситуація</p> <p><b>Result</b> 1 - Це найгрша ситуація 2 - Це сама грша ситуація 3 - Це сама грша ситуація</p>
<p><b>Input sentence</b> Ще 6 ми стільки прочитали сказала Василина.</p> <p><b>Result</b> 1 - Ще 6 ми стільки прочитали, сказала Кирилена. 2 - Ще 6 ми стільки прочитали, сказала Михайліна. 3 - Ще 6 ми так прочитали, сказала Кирилена.</p>	<p><b>Input sentence</b> Ще 6 ми стільки прочитали сказала Василина.</p> <p><b>Result</b> 1 - Ще 6 ми стільки прочитали, сказала Василина. 2 - Ще 6 ми стільки прочитали, – сказала Василина. 3 - Ще 6 ми стільки прочитали сказала Василина.</p>
<p><b>Input sentence</b> Стогла неч красива мов Кармен, червоні й чорні мерла тронди.</p> <p><b>Result</b> 1 - Стогла неч, красива Кармен, червоні й чорні мерла тронди. 2 - Стогла неч, красива Кармен, червоні й білі мерла тронди. 3 - Стогла неч, красива Кармен, білі й чорні мерла тронди.</p>	<p><b>Input sentence</b> Стогла неч красива мов Кармен, червоні й чорні мерла тронди</p> <p><b>Result</b> 1 - Стогла неч красива, мов Кармен, червоні й чорні 2 - Стогла неч красива мов Кармен, червоні й чорні мр 3 - Стогла неч красива, мов Кармен, червоні та чорні мр</p>
<p><b>Input sentence</b> Хоча снігу немає про те надвори мороз пробірас посліни.</p> <p><b>Result</b> 1 - Хоча снігу немає про те надвори мороз пробірас посліни. 2 - Хоча снігу не має про те надвори мороз пробірас посліни. 3 - Хоча снігу не має бути надвори, мороз морози пробірас посліни.</p>	<p><b>Input sentence</b> Хоча снігу немає про те надвори мороз пробірас посліни.</p> <p><b>Result</b> 1 - Хоча снігу немає про те надвори мороз пробірас посліни 2 - Хоча снігу не має про те надвори мороз пробірас посліни 3 - Хоча снігу не має про те надвори мороз пробірас посліни</p>

Рисунок 32 – Варіанти корекції різним NN-моделями

Із наведених прикладів видно, що mT5 краще вдається виправляти граматичні без зміни початкового змісту речення. Однак дана модель має вдвічі більше параметрів, ніж енкодер-декодер, заснований на Roberta (580 мільйонів проти 250), що вимагає значних обчислювальних ресурсів для навчання NN та отримання передбачення. Нейронна мережа M2M100 значно гірше виправляє граматичні помилки, ніж дві попередні моделі (рис. 33)

**Input sentence**  
він за професії архітектор.

**Result**  
1 - Він за професії архітектору.  
2 - Він за професії архітекторе.  
3 - Він за професії архітектор.

**Input sentence**  
Концерт тривав майже півтора години.

**Result**  
1 - Концерт тривав майже півтора години.  
2 - Концерт тривав майже пів року.  
3 - Концерт тривав майже півтори години.

**Input sentence**  
Водій автобуса сказав пасажиром щоб вони оплати проїзд.

**Result**  
1 - Водій автобуса сказав пасажиром, щоб вони оплати проїзд.  
2 - Водій автобуси сказав пасажиром, щоб вони оплати проїзд.  
3 - Водій автобусу сказав пасажиром, щоб вони оплати проїзд.

Рисунок 33 – Робота нейронної мережі M2M100

## 6 ОБГОВОРЕННЯ

Порівняння нейронних мереж за допомогою спеціальних метрик. Відповідно до [54], для оцінки GEC-якості не існує базових, усталених метрик. Із метою оцінки таких систем в загальному використовують метрики MT-якості, однак і вони мають певні недоліки.

BLEU (bilingual evaluation understudy) [55] – це алгоритм для оцінки MT-якості. Оцінки розраховуються для окремих перекладених сегментів (зазвичай речень) шляхом порівняння їх з набором правильних перекладів хорошої якості. Ці бали потім усереднюються по всьому корпусу, щоб отримати оцінку загальної якості перекладу. Оцінка за BLEU завжди є числом від 0 до 1. Це значення вказує, наскільки текст-кандидат подібний до еталонних текстів, а значення, ближчі до 1, означають більш подібні тексти. Значення BLEU засноване на розрахунку кількості спільних N-грам у початковому реченні та його перекладі. Недоліком метрики BLEU є те, що вона заснована на порівняннях стрічок, і, відтак, ігнорується той факт, що деякі граматичні помилки можна виправити кількома способами. Окрім того, результати отримані результати GEC-якості не можна порівняти із результатами певних систем перекладу, оскільки у першому випадку більша кількість n-грамів у реченнях буде співпадати (або співпадати повністю, якщо речення не містить граматичних помилок і не було змінено).

METEOR (Metric for Evaluation of Translation with Explicit ORdering) [56] – ще одна метрика для оцінювання MT-якості. Вона заснована на використанні n-gram та орієнтована на використання статистичної та точної оцінки вихідного тексту. На відміну від метрики BLUE, дана метрика використовує функції співставлення синонімів разом із точною відповідністю слів. Метрика була розроблена, щоб вирішити проблеми, які мала BLUE, а також відтворити хорошу кореляцію з оцінкою експертів на рівні словосполучень або речень.

В результаті запуску метрики на рівні словосполучень кореляція з людським рішенням становила 0.964, тоді кореляція з BLUE становила 0,817 на тому ж наборі вхідних даних. На рівні речень максимальна кореляція з оцінкою експертів була 0,403 [56].

Як і в метриці BLUE, основна одиниця для оцінки – речення, алгоритм спочатку проводить вирівнювання тексту між двома реченнями, рядком еталонного перекладу та рядком вхідного тексту для оцінювання. Дана метрика використовує декілька етапів встановлення відповідності між словами MT й еталонного перекладу для зіставлення двох рядків:

1. Точне встановлення відповідності – визначаються рядки, що є ідентичними в еталонному і машинному перекладі.

2. Встановлення відповідності основ – проводиться стемінг (виділення основи слова) і

визначаються слова з однаковим коренем в еталонному і машинному перекладі.

3. Встановлення відповідності синонімів – визначаються слова, що є синонімами відповідно до WordNet.

Розраховані метрики MT-якості дають змогу лише частково порівняти моделі, оскільки більшість слів і словосполучень у початковому та виправленому реченні співпадають. Найкраще значення як BLEU (0,908), так і METEOR (0,956) отримано для mT5, що співпадає із аналізом прикладів, у якому найбільш точні виправлення помилок без зміни початкового значення речення отримані саме для цієї нейронної мережі. M2M100 має більшу оцінку BLEU (0,847), ніж «Ukrainian Roberta» Encoder-Decoder (0,697), однак, суб'єктивно оцінюючи результати виправлення прикладів, M2M100 значно гірше справляється із цим завданням, ніж дві інші моделі. Для METEOR також M2M100 (0,925) має більшу оцінку, ніж «Ukrainian Roberta» Encoder-Decoder (0,876). Щодо розробки проекту можна зробити такі висновки:

– Для створення якісної ML-моделі для GEC у текстах тих мов, які є складними морфологічно, необхідна велика кількість паралельних або вручну розмічених даних. Ручна анотація даних вимагає багато зусиль професійних лінгвістів, що робить створення корпусів текстів, особливо морфологічно багатих мов, часо- та ресурсозатратним процесом.

– Вирішення завдання автоматичного виявлення та виправлення помилок в україномовних текстах вимагає подальших досліджень у зв'язку з невеликою кількістю робіт, що фокусуються на обробці саме української мови. Окрім того, відповідно до результатів дослідження Погорілого С. Д. і Крамова А. А. [27], методи, які застосовуються для обробки англійської мови, не можуть бути використані для української, оскільки остання є значно складнішою і морфологічно багатшою мовою.

– Поява у 2017 році нової архітектури глибинного навчання Transformer [3], що містить механізм уваги, дозволила значно спростити розробку мовних моделей, що тепер може відбуватися без застосування експертних знань домену, у даному випадку – лінгвістики. Більша частина досліджень застосовує модель Transformer для впровадження нейронного MT тексту із помилками у його правильний варіант. Значним недоліком такого підходу для розробки моделей для різних мов є необхідність у великих корпусах анотованих або паралельних даних. Єдиний (на момент написання роботи) україномовний набір даних [15] містить усього 20 715 речень, такої кількості навчальних зразків не є достатньо для створення автоматичної інтелектуальної MT-системи.

– Застосування перевірки правил та парсингу синтаксичних дерев при розробці GEC-системи в україномовних текстах може бути обумовлене такими двома факторами:

1) відсутністю достатньо великих анотованих / паралельних корпусів української мови для навчання

повністю автоматичних моделей, що засновані на алгоритмах глибинного навчання;

2) недостатньою спроможністю лінгвістичних моделей штучного інтелекту узагальнювати граматичні правила природної мови [37].

– розробка якісної системи перевірки граматичної правильності речень в україномовних текстах вимагатиме поєднання ML-алгоритмів із кількома різними типами методів, зокрема і застосування експертних знань з комп'ютерної лінгвістики.

– Відповідно до результатів досліджень, для отримання найкращої GEC-точності за допомогою нейронної мережі, потрібно використовувати попередньо навчені MT-моделі, що підтримують українську мову.

– Найкраще значення як BLEU, так і METEOR було отримано для MT-моделі mT5, результати співпадають із аналізом власних прикладів, у якому найбільш точні виправлення помилок без зміни початкового значення речення були отримані саме для цієї нейронної мережі. M2M100 має більшу оцінку BLEU, ніж «Ukrainian Roberta» Encoder-Decoder, однак, суб'єктивно оцінюючи результати виправлення власних прикладів, M2M100 значно гірше справляється із цим завданням, ніж дві інші моделі. З метою економії обчислювальних ресурсів можливим також є застосування попередньо навченої нейронної мережі типу BERT, використовуючи її як у якості енкодера, так і декодера. Така нейронна мережа має вдвічі менше параметрів, ніж інші попередньо навчені MT-моделі, і показує задовільні GEC-результати.

### ВИСНОВКИ

Результатом роботи є розроблена ML-модель для GEC в україномовних текстах. Ми також пропонуємо універсальну схему розробки GEC-системи для різних мов. Відповідно до отриманих результатів, нейронна мережа має здатність виправляти прості речення, написані українською, однак розробка повноцінної системи вимагатиме застосування перевірки орфографії за допомогою словників і перевірки правил, як простих, так і заснованих на результаті парсингу залежностей або інших ознак.

З-поміж трьох моделей, найкращі показники має попередньо навчена модель нейронного перекладу mT5. Найкраще значення як BLEU (0,908), так і METEOR (0,956) отримано для mT5, що співпадає із аналізом прикладів, у якому найбільш точні виправлення помилок без зміни початкового значення речення отримані саме для цієї нейронної мережі. M2M100 має більшу оцінку BLEU (0,847), ніж «Ukrainian Roberta» Encoder-Decoder (0,697), однак, суб'єктивно оцінюючи результати виправлення прикладів, M2M100 значно гірше справляється із цим завданням, ніж дві інші моделі. Для METEOR також M2M100 (0,925) має більшу оцінку, ніж «Ukrainian Roberta» Encoder-Decoder (0,876).

З метою економії обчислювальних ресурсів можливим також є застосування попередньо навченої нейронної мережі типу BERT, використовуючи її як у якості енкодера, так і декодера. Така нейронна мережа має вдвічі менше параметрів, ніж інші попередньо навчені MT-моделі, і показує задовільні GEC-результати.

### ПОДЯКИ

Роботу виконано в рамках держбюджетної теми «Методи та засоби функціонування систем підтримки прийняття рішень на основі онтологій» (ID:839 2017-05-15 09:20:01 (2459-315)). Дослідження провадилося в межах спільних наукових досліджень кафедри інформаційних систем та мереж НУ «Львівська політехніка» на тему «Дослідження, розроблення і впровадження інтелектуальних розподілених інформаційних технологій та систем на основі ресурсів баз даних, сховищ даних, просторів даних та знань з метою прискорення процесів формування сучасного інформаційного суспільства». Наукові дослідження провадилися також в рамках ініціативної тематики досліджень кафедри ІСМ НУ «Львівська політехніка» на тему «Розроблення інтелектуальних розподілених систем на основі онтологічного підходу з метою інтеграції інформаційних ресурсів».

### ЛІТЕРАТУРА

1. Naghshnejad M. Recent Trends in the Use of Deep Learning Models for Grammar Error Handling / M. Naghshnejad, T. Joshi, V. N. Nair // ArXiv. – 2020. DOI: 10.48550/arXiv.2009.02358
2. Automated Grammatical Error Detection for Language Learners, Second Edition / [C. Leacock, M. Chodorow, M. Gamon, J. Tetreault] // Synthesis Lectures on Human Language Technologies. – 2014, Berlin : Springer. – 154 p. DOI: 10.1007/978-3-031-02153-4
3. Attention Is All You Need / [A. Vaswani, N. Shazeer, N. Parmar et al.] // ArXiv. – 2017. DOI: 10.48550/arXiv.1706.03762
4. Transformers: State-of-the-Art Natural Language Processing / [T. Wolf, L. Debut, V. Sanh et al.] // EMNLP 2020: Conference on Empirical Methods in Natural Language Processing: System Demonstrations, Online, ALC Anthology, Oct. 2020 : proceedings. – P. 38–45. DOI: 10.18653/v1/2020.emnlp-demos.6
5. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding / [J. Devlin, M.-W. Chang, K. Lee, K. Toutanova] // Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies Minneapolis, Minnesota, ALC Anthology, June 2019 : proceedings. – P. 4171–4186. DOI: 10.18653/v1/N19-1423
6. Rozovskaya A. Grammar Error Correction in Morphologically Rich Languages: The Case of Russian / A. Rozovskaya, D. Roth // Transactions of the Association for Computational Linguistics. – 2019. – Vol. 7. – P. 1–17. DOI: 10.1162/tacl\_a\_00251
7. Bick E. DanProof: Pedagogical Spell and Grammar Checking for Danish / E. Bick // International Conference Recent Advances in Natural Language Processing, Hissar, Bulgaria, Sep. 2015 : proceedings. – P. 55–62.



8. Design and construction of the Greek grammar checker / [P. Gakis, C. T. Panagiotakopoulos, K. N. Sgarbas et al.] // *Digital Scholarship in the Humanities*. – 2017. – Vol. 32. – P. 554–576. DOI: 10.1093/lc/fqw025
9. Deksne D. A New Phase in the Development of a Grammar Checker for Latvian / D. Deksne // *Human Language Technologies – The Baltic Perspective*. – IOS Press, 2016. – P. 147–152. DOI: 10.3233/978-1-61499-701-6-147
10. Sorokin A. Spelling Correction for Morphologically Rich Language: a Case Study of Russian / A. Sorokin // *6th Workshop on Balto-Slavic Natural Language Processing, Valencia, Spain, Apr. 2017* : proceedings. – P. 45–53. DOI: 10.18653/v1/W17-1408.
11. Gill M. S. A Grammar Checking System for Punjabi / M. S. Gill, G. S. Lehal // *Coling 2008: Companion volume: Demonstrations, Manchester, UK, Aug. 2008* : proceedings. – P. 149–152.
12. Go M. P. Developing an Unsupervised Grammar Checker for Filipino Using Hybrid N-grams as Grammar Rules / M. P. Go, A. Borra // *PACLIC: 30th Pacific Asia Conference on Language, Information and Computation: Oral Papers, Seoul, South Korea, Oct. 2016* : proceedings. – P. 105–113.
13. Shaaan K. F. Arabic GramCheck: a grammar checker for Arabic / K. F. Shaaan // *Software: Practice and Experience*. – 2005. – Vol. 35(7). – P. 643–665. DOI: 10.1002/spe.653
14. A Comprehensive Survey of Grammar Error Correction / [Y. Wang, Y. Wang, J. Liu, Z. Liu] // *ArXiv*. – 2020. DOI: 10.48550/arXiv.2005.06600
15. Syvokon O. UA-GEC: Grammatical Error Correction and Fluency Corpus for the Ukrainian Language / O. Syvokon, O. Nahorna // *ArXiv*. – 2021. DOI: 10.48550/arXiv.2103.16997
16. Lardinois F. Grammarly goes beyond grammar / F. Lardinois // *TechCrunch*. – 2019. – Access mode: <https://techcrunch.com/2019/07/16/grammarly-goes-beyond-grammar/>.
17. Lardinois F. Grammarly gets a tone detector to keep you out of email trouble / F. Lardinois // *TechCrunch*. – 2019. – Access mode: <https://techcrunch.com/2019/09/24/grammarly-gets-a-tone-detector-to-keep-you-out-of-email-trouble/>.
18. Grammarly Inc. About Us / Grammarly Inc. – Access mode: <https://www.grammarly.com/about>.
19. Grammarly Inc. Does Grammarly support languages other than English? / Grammarly Inc. – Access mode: <https://support.grammarly.com/hc/en-us/articles/115000090971-Does-Grammarly-support-languages-other-than-English->.
20. LanguageTool. Languages / LanguageTool. – Access mode: <https://dev.languagetool.org/languages>.
21. LanguageTool. Error Rules for LanguageTool / LanguageTool Community. – Access mode: [https://community.languagetool.org/rule/list?offset=0&max=10&lang=uk&filter=&categoryFilter=&\\_action\\_list=%D0%A4%D1%96%D0%BB%D1%8C%D1%82%D1%80](https://community.languagetool.org/rule/list?offset=0&max=10&lang=uk&filter=&categoryFilter=&_action_list=%D0%A4%D1%96%D0%BB%D1%8C%D1%82%D1%80).
22. LanguageTool. About / LanguageTool. – Access mode: <https://languagetool.org/about>.
23. Jayanthi S. NeuSpell: A Neural Spelling Correction Toolkit / S. Jayanthi, D. Pruthi, G. Neubig // *ArXiv*. – 2020. DOI: 10.48550/arXiv.2010.11085
24. Hunspell, Github. – 2021. – Access mode: <https://github.com/hunspell/hunspell>.
25. Korobov M. Morphological Analyzer and Generator for Russian and Ukrainian Languages / M. Korobov // *ArXiv*. – 2015. DOI: 10.48550/arXiv.1503.07283
26. Tmienova N. System of Intellectual Ukrainian Language Processing / N. Tmienova, B. Sus // *ITS: the XIX International Conference on Information Technologies and Security, Kyiv, Ukraine, Nov. 28, 2019* : proceedings. – P. 199–209.
27. Pogorilyy S. Method of noun phrase detection in Ukrainian texts / S. Pogorilyy, A. A. Kramov // *ArXiv*. – 2020. DOI: 10.48550/arXiv.2010.11548
28. Глибовець А. Алгоритм токенизації та стемінгу для текстів українською мовою / А. Глибовець, В. Точицький // *Наукові записки НаУКМА. Комп'ютерні науки*. – 2017. – Т. 198. – С. 4–8.
29. Hao S. A Research on Online Grammar Checker System Based on Neural Network Model / S. Hao, G. Hao // *Journal of Physics*. – 2020. – Vol. 1651. – P. 1–8. DOI: 10.1088/1742-6596/1651/1/012135
30. Батюк Т. М. Технологія соціалізації особистостей за спільними інтересами на основі методів машинного навчання та seo-технологій / Т. М. Батюк, В. А. Висоцька // *Радіоелектроніка, інформатика, управління*. – 2022. – № 2 (61). – С. 53–68. DOI: 10.15588/1607-3274-2022-2-6
31. The CoNLL-2014 Shared Task on Grammatical Error Correction / [H. T. Ng, S. M. Wu, T. Briscoe, et al.] // *Conference on Computational Natural Language Learning: Shared Task, Baltimore, Maryland, Jun. 2014* : proceedings. – P. 1–14. DOI: 10.3115/v1/W14-1701
32. Chollampatt S. A Multilayer Convolutional Encoder-Decoder Neural Network for Grammatical Error Correction / S. Chollampatt, H. T. Ng // *ArXiv*. – 2018. DOI: 10.48550/arXiv.1801.08831
33. GECToR – Grammatical Error Correction: Tag, Not Rewrite / [K. Omelianchuk, V. Atrasevych, A. N. Chernodub, O. Skurzhanyskiy] // *ArXiv*. – 2020. DOI: 10.48550/arXiv.2005.12592
34. The University of Illinois System in the CoNLL-2013 Shared Task / [A. Rozovskaya, K.-W. Chang, M. Sammons, D. Roth] // *Conference on Computational Natural Language Learning: Shared Task, Sofia, Bulgaria, Aug. 2013* : proceedings. – P. 13–19.
35. The CoNLL-2013 Shared Task on Grammatical Error Correction / [H. T. Ng, S. M. Wu, Y. Wu et al.] // *Conference on Computational Natural Language Learning: Shared Task, Sofia, Bulgaria, Aug. 2013* : proceedings. – P. 1–12.
36. A Simple Recipe for Multilingual Grammatical Error Correction / [S. Rothe, J. Mallinson, E. Malmi et al.] // *ArXiv*. – 2021. DOI: 10.48550/arXiv.2106.03830
37. Mita M. Do Grammatical Error Correction Models Realize Grammatical Generalization? / M. Mita, H. Yanaka // *ArXiv*. – 2021. DOI: 10.48550/arXiv.2106.03031
38. A Unified Strategy for Multilingual Grammatical Error Correction with Pre-trained Cross-Lingual Language Model / [X. Sun, T. Ge, S. Ma et al.] // *ArXiv*. – 2022. DOI: 10.48550/arXiv.2201.10707
39. Yasunaga M. LM-Critic: Language Models for Unsupervised Grammatical Error Correction / M. Yasunaga, J. Leskovec, and P. Liang // *ArXiv*. – 2021. DOI: 10.48550/arXiv.2109.06822
40. A Neural Grammatical Error Correction System Built on Better Pre-training and Sequential Transfer Learning / [Y. J. Choe, J. Ham, K. Park, Y. Yoon] // *Workshop on Innovative*

- Use of NLP for Building Educational Applications, Florence, Italy, Aug. 2019 : proceedings. – P. 213–227. DOI: 10.18653/v1/W19-4423
41. Wan Z. A Syntax-Guided Grammatical Error Correction Model with Dependency Tree Correction / Z. Wan, X. Wan // ArXiv. – 2021. DOI: 10.48550/arXiv.2111.03294
42. Neural Language Correction with Character-Based Attention / [Z. Xie, A. Avati, N. Arivazhagan et al.] // ArXiv. – 2016. DOI: 10.48550/arXiv.1603.09727
43. Parnow K. Grammatical Error Correction as GAN-like Sequence Labeling / K. Parnow, Z. Li, H. Zhao // ArXiv. – 2021. DOI: 10.48550/arXiv.2105.14209
44. Wang X. Research and Implementation of English Grammar Check and Error Correction Based on Deep Learning / X. Wang, W. Zhong // Scientific Programming. – 2022. – Vol. 2022. – Article ID 4082082. DOI: 10.1155/2022/4082082
45. Grammatical Error Correction in Low-Resource Scenarios / J. Náplava, M. Straka // W-NUT: 5th Workshop on Noisy User-generated Text, Hong Kong, China, Nov. 2019 : proceedings. – P. 346–356. DOI: 10.18653/v1/D19-5545.
46. Improving Grammatical Error Correction with Machine Translation Pairs / [W. Zhou, T. Ge, C. Mu et al.] // ArXiv. – 2020. DOI: 10.48550/arXiv.1911.02825
47. Raheja V. Adversarial Grammatical Error Correction / V. Raheja, D. Alikaniotis // EMNLP: Findings of the Association for Computational Linguistics, Online, Nov. 2020 : proceedings. – P. 3075–3087. DOI: 10.18653/v1/2020.findings-emnlp.275
48. Radchenko V. Ukrainian Roberta / V. Radchenko. Access mode: <https://github.com/youscan/language-models>.
49. mT5: A Massively Multilingual Pre-trained Text-to-Text Transformer / [L. Xue, N. Constant, A. Roberts et al.] // Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Online, Jun. 2021 : proceedings. – P. 483–498. DOI: 10.18653/v1/2021.naacl-main.41.
50. Beyond English-Centric Multilingual Machine Translation / [A. Fan, S. Bhosale, H. Schwenk et al.] // ArXiv. – 2020. DOI: 10.48550/arXiv.2010.11125
51. Multilingual Translation with Extensible Multilingual Pretraining and Finetuning / [Y. Tang, C. Tran, Xian Li et al.] // ArXiv. – 2020. DOI: 10.48550/arXiv.2008.00401
52. Platen von Patrick Leveraging Pre-trained Language Model Checkpoints for Encoder-Decoder Models / Platen von Patrick. – Access: <https://huggingface.co/blog/warm-starting-encoder-decoder>.
53. Rothe S. Leveraging Pre-trained Checkpoints for Sequence Generation Tasks / S. Rothe, S. Narayan, A. Severyn // ArXiv. – 2019. DOI: 10.48550/arXiv.1907.12461
54. Ground truth for grammatical error correction metrics / [C. Napoles, K. Sakaguchi, M. Post, J. Tetreault] // 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, Beijing, China, July 2015 : proceedings. – P. 588–593. DOI: 10.3115/v1/P15-2097
55. Bleu: a Method for Automatic Evaluation of Machine Translation / [K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu] // 40th Annual Meeting of the Association for Computational Linguistics, Philadelphia, Pennsylvania, USA, July 2002 : proceedings. – P. 311–318. DOI: 10.3115/1073083.1073135.
56. Banerjee S. METEOR: An Automatic Metric for MT Evaluation with Improved Correlation with Human Judgments / S. Banerjee, A. Lavie // ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization, Michigan, Jun. 2005 : proceedings. – P. 65–72.
57. Platen von Patrick Encoder-Decoder models don't need costly pre-training to yield state-of-the-art results on seq2seq tasks / Platen von Patrick. – Access mode: <https://twitter.com/patrickplaten/status/1325844244095971328>.

Accepted 19.09.2022.  
Received 03.10.2022.

UDC 004.9

#### TECHNOLOGY FOR GRAMMATICAL ERRORS CORRECTION IN UKRAINIAN TEXT CONTENT BASED ON MACHINE LEARNING METHODS

**Kholodna N.** – PhD student of Information Systems and Networks Department, Lviv Polytechnic National University, Lviv, Ukraine.

**Vysotska V.** – PhD, Associate Professor of Information Systems and Networks Department, Lviv Polytechnic National University, Lviv, Ukraine.

#### ABSTRACT

**Context.** Most research in grammatical and stylistic error correction focuses on error correction in English-language textual content. Thanks to the availability of large data sets, a significant increase in the accuracy of English grammar correction has been achieved. Unfortunately, there are few studies on other languages. Systems for the English language are constantly developing and currently actively use machine learning methods: classification (sequence tagging) and machine translation. A large amount of parallel or manually labelled data is required to build a high-quality machine learning model for correcting grammatical/stylistic errors in the texts of those morphologically complex languages. Manual data annotation requires a lot of effort by professional linguists, which makes the creation of text corpora, especially in morphologically rich languages, mainly Ukrainian, a time- and resource-consuming process.

**Objective** of the study is to develop a technology for correcting errors in Ukrainian-language texts based on machine learning methods using a small set of annotated parallel data.

**Method.** For this study, machine learning algorithms were selected when developing a system for correcting errors in Ukrainian-language texts using an optimal pipeline, including pre-processing and selecting text content and generating features in small annotated data corpora. The neural network's use with a new architecture, a review of state-of-the-art methods, and a comparison of different pipeline stages will make it possible to determine such a combination of them, allowing a high-quality error correction model in Ukrainian-language texts.

**Results.** A machine learning model for error correction in Ukrainian-language texts has been developed. A universal scheme for creating an error correction system for different languages is proposed. According to the results, the neural network can correct simple sentences written in Ukrainian. However, creating a full-fledged system will require spell-checking using dictionaries and checking rules, both simple and based on the result of parsing dependencies or other features. The pre-trained neural translation

© Холодна Н. М., Висоцька В. А., 2023  
DOI 10.15588/1607-3274-2023-1-12

model mT5 has the best performance among the three models. To save computing resources, it is also possible to use a pre-trained BERT-type neural network as an encoder and a decoder. Such a neural network has half the number of parameters as other pre-trained machine translation models and shows satisfactory results in correcting grammatical and stylistic errors.

**Conclusions.** The created model shows excellent classification results on test data. The calculated machine translation quality metrics allow only a partial comparison of the models since most of the words and phrases in the original and corrected sentences are the same. The best value for both BLEU (0.908) and METEOR (0.956) is obtained for mT5, which is consistent with the case study in which the most accurate error corrections without changing the initial value of the sentence are obtained for such a neural network. The M2M100 has a higher BLEU score (0.847) than the “Ukrainian Roberta” Encoder-Decoder (0.697). However, subjectively evaluating the results of the correction of examples, the M2M100 does a much worse job than the other two models. For METEOR, M2M100 (0.925) also has a higher score than the “Ukrainian Roberta” Encoder-Decoder (0.876).

**KEYWORDS:** NLP, text pre-processing, error correction, grammatical error correction, machine learning, deep learning, text analysis, text classification, neural network.

#### REFERENCES

1. Naghshnejad M., Joshi T., Nair V. N. Recent Trends in the Use of Deep Learning Models for Grammar Error Handling, *ArXiv*, 2020. DOI: 10.48550/arXiv.2009.02358
2. Leacock C., Chodorow M., Gamon M., Tetreault J. Automated Grammatical Error Detection for Language Learners, Second Edition, *Synthesis Lectures on Human Language Technologies*, 2014. Berlin, Springer, 154 p. DOI: 10.1007/978-3-031-02153-4
3. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser L., Polosukhin I. Attention Is All You Need, *ArXiv*, 2017. DOI: 10.48550/arXiv.1706.03762
4. Wolf T., Debut L., Sanh V., Chaumond J., Delangue C., Moi A., Cistac P., Rault T., Louf R., Funtowicz M., Davison J., Shleifer S., Platen P. v., Ma C., Jernite Y., Plu J., Xu C., Scio T. L., Gugger S., Drame M., Lhoest Q., Rush A. Transformers: State-of-the-Art Natural Language Processing, *EMNLP 2020: Conference on Empirical Methods in Natural Language Processing: System Demonstrations, Online, ALC Anthology, Oct. 2020 : proceedings*, pp. 38–45. DOI: 10.18653/v1/2020.emnlp-demos.6
5. Devlin J., Chang M.-W., Lee K., Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies Minneapolis*. Minnesota, ALC Anthology, June 2019 : proceedings, pp. 4171–4186. DOI: 10.18653/v1/N19-1423
6. Rozovskaya A., Roth D. Grammar Error Correction in Morphologically Rich Languages: The Case of Russian, *Transactions of the Association for Computational Linguistics*, 2019, Vol. 7, pp. 1–17. DOI: 10.1162/tacl\_a\_00251
7. Bick E. DanProof: Pedagogical Spell and Grammar Checking for Danish, *International Conference Recent Advances in Natural Language Processing*. Hissar, Bulgaria, Sep. 2015, proceedings, pp. 55–62.
8. Gakis P., Panagiotakopoulos C. T., Sgarbas K. N., Tsalidis C., Verykios V. S., Design and construction of the Greek grammar checker, *Digital Scholarship in the Humanities*, 2017, Vol. 32, pp. 554–576. DOI: 10.1093/lc/fqw025
9. Deksnė D. A New Phase in the Development of a Grammar Checker for Latvian. Human Language Technologies, The Baltic Perspective, IOS Press, 2016, pp. 147–152. DOI: 10.3233/978-1-61499-701-6-147
10. Sorokin A. Spelling Correction for Morphologically Rich Language: a Case Study of Russian, *6th Workshop on Balto-Slavic Natural Language Processing, Valencia, Spain, Apr. 2017, proceedings*, pp. 45–53. DOI: 10.18653/v1/W17-1408.
11. Gill M. S., Lehal G. S. A Grammar Checking System for Punjabi, *Coling 2008: Companion volume: Demonstrations, Manchester, UK, Aug. 2008 : proceedings*, pp. 149–152.
12. Go M. P., Borra A. Developing an Unsupervised Grammar Checker for Filipino Using Hybrid N-grams as Grammar Rules, *PACLIC: 30th Pacific Asia Conference on Language, Information and Computation, Oral Papers*. Seoul, South Korea, Oct. 2016 : proceedings, pp. 105–113.
13. Shaalan K. F. Arabic GramCheck: a grammar checker for Arabic, *Software: Practice and Experience*, 2005, Vol. 35(7), pp. 643–665. DOI: 10.1002/spe.653
14. Wang Y., Wang Y., Liu J., Liu Z. A Comprehensive Survey of Grammar Error Correction, *ArXiv*, 2020. DOI: 10.48550/arXiv.2005.06600
15. Syvokon O., Nahorna O. UA-GEC: Grammatical Error Correction and Fluency Corpus for the Ukrainian Language *ArXiv*, 2021. DOI: 10.48550/arXiv.2103.16997
16. Lardinois F. Grammarly goes beyond grammar, *TechCrunch*, 2019. Access mode: <https://techcrunch.com/2019/07/16/grammarly-goes-beyond-grammar/>.
17. Lardinois F. Grammarly gets a tone detector to keep you out of email trouble, *TechCrunch*, 2019. Access mode: <https://techcrunch.com/2019/09/24/grammarly-gets-a-tone-detector-to-keep-you-out-of-email-trouble/>.
18. Grammarly Inc. About Us / Grammarly Inc. Access mode: <https://www.grammarly.com/about>.
19. Grammarly Inc. Does Grammarly support languages other than English? / Grammarly Inc. Access mode: <https://support.grammarly.com/hc/en-us/articles/115000090971-Does-Grammarly-support-languages-other-than-English->.
20. LanguageTool. Languages, LanguageTool. Access mode: <https://dev.languagetool.org/languages>.
21. LanguageTool. Error Rules for LanguageTool / LanguageTool Community. Access mode: [https://community.languagetool.org/rule/list?offset=0&max=10&lang=uk&filter=&categoryFilter=&\\_action\\_list=%D0%A4%D1%96%D0%BB%D1%8C%D1%82%D1%80](https://community.languagetool.org/rule/list?offset=0&max=10&lang=uk&filter=&categoryFilter=&_action_list=%D0%A4%D1%96%D0%BB%D1%8C%D1%82%D1%80).
22. LanguageTool. About / LanguageTool. Access mode: <https://languagetool.org/about>.
23. Jayanthi S., Pruthi D., Neubig G. NeuSpell: A Neural Spelling Correction Toolkit, *ArXiv*, 2020. DOI: 10.48550/arXiv.2010.11085
24. Hunspell, Github, 2021, Access mode: <https://github.com/hunspell/hunspell>.
25. Korobov M. Morphological Analyzer and Generator for Russian and Ukrainian Languages, *ArXiv*, 2015. DOI: 10.48550/arXiv.1503.07283
26. Tmienova N., Sus B. System of Intellectual Ukrainian Language Processing, *ITS: the XIX International Conference on Information Technologies and Security, Kyiv, Ukraine, Nov. 28, 2019 : proceedings*, pp. 199–209.

27. Pogorilyy S., Kramov A. A. Method of noun phrase detection in Ukrainian texts, *ArXiv*, 2020. DOI: 10.48550/arXiv.2010.11548
28. Hlybovets A., Tochytyskiy V. Algorithm of tokenization and stemming for texts in the Ukrainian language, *Scientific notes of NaUKMA. Computer Science*, 2017, Vol. 198, pp. 4–8.
29. Hao S., Hao G. A Research on Online Grammar Checker System Based on Neural Network Model, *Journal of Physics*, 2020, Vol. 1651, pp. 1–8. DOI: 10.1088/1742-6596/1651/1/012135
30. Batiuk T. M., Vysotska V. Technology for Personalities Socialization by Common Interests Based on Machine Learning Methods And SEO-Technologies, *Radio Electronics, Computer Science, Control*, 2022, Vol. 2 (61), pp. 53–68. DOI: 10.15588/1607-3274-2022-2-6
31. Ng H. T., Wu S. M., Briscoe T., Hadiwinoto C., Susanto R. H., Bryant C. The CoNLL-2014 Shared Task on Grammatical Error Correction, *Conference on Computational Natural Language Learning: Shared Task*. Baltimore, Maryland, Jun. 2014, proceedings, pp. 1–14. DOI: 10.3115/v1/W14-1701
32. Chollampatt S., Ng H. T. A Multilayer Convolutional Encoder-Decoder Neural Network for Grammatical Error Correction, *ArXiv*, 2018. DOI: 10.48550/arXiv.1801.08831
33. Omelianchuk K., Atrasevych V., Chernodub A. N., Skurzhanyskiy O. GECToR – Grammatical Error Correction: Tag, Not Rewrite, *ArXiv*, 2020. DOI: 10.48550/arXiv.2005.12592
34. Rozovskaya A., Chang K.-W., Sammons M., Roth D. The University of Illinois System in the CoNLL-2013 Shared Task, *Conference on Computational Natural Language Learning: Shared Task*. Sofia, Bulgaria, Aug. 2013, proceedings, pp. 13–19.
35. Ng H. T., Wu S. M., Wu Y., Hadiwinoto C., Tetreault J. The CoNLL-2013 Shared Task on Grammatical Error Correction, *Conference on Computational Natural Language Learning: Shared Task*. Sofia, Bulgaria, Aug. 2013, proceedings, pp. 1–12.
36. Rothe S., Mallinson J., Malmi E., Krause S., Severyn A. A Simple Recipe for Multilingual Grammatical Error Correction, *ArXiv*, 2021. DOI: 10.48550/arXiv.2106.03830
37. Mita M., Yanaka H. Do Grammatical Error Correction Models Realize Grammatical Generalization?, *ArXiv*, 2021. DOI: 10.48550/arXiv.2106.03031
38. Sun X., Ge T., Ma S., Li J., Wei F., and Wang H. A Unified Strategy for Multilingual Grammatical Error Correction with Pre-trained Cross-Lingual Language Model, *ArXiv*, 2022. DOI: 10.48550/arXiv.2201.10707
39. Yasunaga M., Leskovec J., and Liang P. LM-Critic: Language Models for Unsupervised Grammatical Error Correction, *ArXiv*, 2021. DOI: 10.48550/arXiv.2109.06822
40. Choe Y. J., Ham J., Park K., Yoon Y. A Neural Grammatical Error Correction System Built on Better Pre-training and Sequential Transfer Learning, *Workshop on Innovative Use of NLP for Building Educational Applications*, Florence. Italy, Aug. 2019 : proceedings, pp. 213–227. DOI: 10.18653/v1/W19-4423
41. Wan Z., Wan X. A Syntax-Guided Grammatical Error Correction Model with Dependency Tree Correction, *ArXiv*, 2021. DOI: 10.48550/arXiv.2111.03294
42. Xie Z., Avati A., Arivazhagan N., Jurafsky D., Ng A. Neural Language Correction with Character-Based Attention, *ArXiv*, 2016. DOI: 10.48550/arXiv.1603.09727
43. Parnow K., Li Z., Zhao H. Grammatical Error Correction as GAN-like Sequence Labeling, *ArXiv*, 2021. DOI: 10.48550/arXiv.2105.14209
44. Wang X., Zhong W. Research and Implementation of English Grammar Check and Error Correction Based on Deep Learning, *Scientific Programming*, 2022, Vol. 2022, Article ID 4082082. DOI: 10.1155/2022/4082082
45. Náplava J., Straka M. Grammatical Error Correction in Low-Resource Scenarios, *W-NUT: 5th Workshop on Noisy User-generated Text*. Hong Kong, China, Nov. 2019 : proceedings, pp. 346–356. DOI: 10.18653/v1/D19-5545.
46. Zhou W., Ge T., Mu C., Xu K., Wei F., Zhou M. Improving Grammatical Error Correction with Machine Translation Pairs, *ArXiv*, 2020. DOI: 10.48550/arXiv.1911.02825
47. Raheja V., Alikaniotis D. Adversarial Grammatical Error Correction, *EMNLP: Findings of the Association for Computational Linguistics, Online*, Nov. 2020 : proceedings, pp. 3075–3087. DOI: 10.18653/v1/2020.findings-emnlp.275
48. Radchenko V. Ukrainian Roberta. Access mode: <https://github.com/youscan/language-models>.
49. Xue L., Constant N., Roberts A., Kale M., Al-Rfou R., Siddhant A., Barua A., Raffel C. mT5: A Massively Multilingual Pre-trained Text-to-Text Transformer, Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Online, Jun. 2021 : proceedings, pp. 483–498. DOI: 10.18653/v1/2021.naacl-main.41.
50. Fan A., Bhosale S., Schwenk H., Ma Z., El-Kishky A., Goyal S., Baines M., Celebi O., Wenzek G., Chaudhary V., Goyal N., Birch T., Liptchinsky V., Edunov S., Grave E., Auli M., Joulin A. Beyond English-Centric Multilingual Machine Translation, *ArXiv*, 2020. DOI: 10.48550/arXiv.2010.11125
51. Tang Y., Tran C., Li Xian, Chen P.-J., Goyal N., Chaudhary V., Gu J., Fan A. Multilingual Translation with Extensible Multilingual Pretraining and Finetuning, *ArXiv*, 2020. DOI: 10.48550/arXiv.2008.00401
52. Platen von Patrick Leveraging Pre-trained Language Model Checkpoints for Encoder-Decoder Models. Access: <https://huggingface.co/blog/warm-starting-encoder-decoder>.
53. Rothe S., Narayan S., Severyn A. Leveraging Pre-trained Checkpoints for Sequence Generation Tasks, *ArXiv*, 2019. DOI: 10.48550/arXiv.1907.12461
54. Napoles C., Sakaguchi K., Post M., Tetreault J. Ground truth for grammatical error correction metrics, *53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*. Beijing, China, July 2015 : proceedings, pp. 588–593. DOI: 10.3115/v1/P15-2097
55. Papineni K., Roukos S., Ward T., and Zhu W.-J. Bleu: a Method for Automatic Evaluation of Machine Translation, *40th Annual Meeting of the Association for Computational Linguistics, Philadelphia, Pennsylvania, USA, July 2002 : proceedings*, pp. 311–318. DOI: 10.3115/1073083.1073135.
56. Banerjee S., Lavie A. METEOR: An Automatic Metric for MT Evaluation with Improved Correlation with Human Judgments, *ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization*. Michigan, Jun. 2005 : proceedings, pp. 65–72.
57. Platen von Patrick Encoder-Decoder models don't need costly pre-training to yield state-of-the-art results on seq2seq tasks. Access mode: <https://twitter.com/patrickplaten/status/1325844244095971328>.

# УПРАВЛІННЯ У ТЕХНІЧНИХ СИСТЕМАХ

## CONTROL IN TECHNICAL SYSTEMS

UDK 519.853: 658.52

### METHOD OF ROUTING A GROUP OF MOBILE ROBOTS IN A FIXED NETWORK FOR SEARCHING THE MISSING OBJECTS IN A TECHNOLOGICAL DISASTER ZONE

**Batsamut V. M.** – Dr. Sc., Professor, The Deputy Head of the Scientific Research Center of the National Academy of the National Guard of Ukraine, Kharkiv, Ukraine.

**Hodlevsky S. O.** – The Researcher of the Scientific Research Center of the National Academy of the National Guard of Ukraine, Kharkiv, Ukraine.

#### ABSTRACT

**Context.** The relevance of the article is determined by the need for further development of models of collective behavior of systems with multi-agent structure construction endowed with intelligence that ensures synchronization of the joint efforts of various agents while achieving the goals set for the system. The method proposed in the article solves the problem of competition between different agents of a multi-agent system, which is important while performing search, rescue, and monitoring tasks in crisis areas of various origins.

**The purpose of the study** is to develop a method for determining the sufficient population of a multi-agent system and the optimal routes of movement of its individual elements in a stationary network for the most complete examination of a technological disaster zone (any given zone based on a certain transport network).

**Method.** We implemented the concept of a dynamic programming to search for all possible edge-simple longest paths connecting the directed subsets of vertices-sources and vertices-sinks in the structure of the model weighted directed graph. To this end, the modified Dijkstra method was applied. The modification comprises representing the weights of the arcs of the modeling directed graph with the negative values, which are further used in calculations according to the Dijkstra method. After finding the next edge-simple longest path, the arcs that make up it are fixed in the memory of the computer system (in the route plan) and removed from the graph structure, and the process is iteratively repeated. The search for paths takes place as long as the transitive closure between the vertices that are part of the specified subsets of source vertices and sink vertices is preserved. The developed method makes it possible to find such a set of traffic routes for the elements of the multi-agent system, which maximizes the area examined by them in a technological disaster zone (or the number of checked objects on the traffic routes) in one "wave" of the search and distributes the elements of a multi-agent system by routes that do not have common areas. A derivative of the application of the developed method is the determination of a sufficient population of a multi-agent system for effective search activities within the defined zone.

**Results.** 1) A method of routing a group of mobile robots in a stationary network for searching the missing objects in a technological disaster zone has been developed. 2) The working expression of the Dijkstra method for searching in the structure of a network object (in the structure of a model graph) for the longest paths has been formalized. 3) We have suggested a set of indicators for a comprehensive evaluation of route plans of a multi-agent system. 4) The method has been verified on test problems.

**Conclusions.** Theoretical studies and several experiments confirm the efficiency of the developed method. The solutions made using the developed method are accurate, which allows recommending it for practical use in determining in an automated mode route plans for multi-agent systems, as well as the required number of agents in such systems to perform the required amount of search tasks in a particular crisis area.

**KEYWORDS:** multi-agent system, group of mobile robots, routing, network object, weighted undirected (directed) graph, extreme paths, optimization criterion, method.

#### ABBREVIATIONS

MAS is a multi-agent system;  
TDZ is a technological disaster zone.

#### NOMENCLATURE

$G$  is a undirected weighted graph simulating the transport network in a technological disaster zone;

$\vec{G}$  is a directed weighted graph simulating the transport network in a technological disaster zone;

$P$  is a set of graph vertices simulating turning points (intersections) of the transport network within a technological disaster zone;

$A$  is a subset of vertices-sources such that  $A \subset P$ ;

$B$  is a subset of vertices-sinks such that  $B \subset P$ ;

$E$  is a set of edges (arcs) of the graph modeling paths (communications) inside a technological disaster zone;

$w(e_{ij})$  is a weighting coefficient of some edge (arc)  $e_{ij}$ ;

*Direct* is a general search direction;

$k_{wvs}$  is a number of searching waves;

$M_i$  is the longest edge-simple path between some vertices  $a_x$  and  $b_y$  from sets  $A$  and  $B$  respectively;

$\vec{G}'$ ,  $\vec{G}''$ , ...,  $\vec{G}^n$  is a sequence of substructures arising because of splitting the initial graph  $\vec{G}$ ;

$n$  is the number of found longest edge-simple paths between subsets  $A$  and  $B$  respectively;

$L_i$  is the length of  $i$ -th edge-simple longest path;

$i$  is the number of the edge-simple longest paths;

$L_{total}$  is the total length of the transport network in the TDZ;

$L_{PL}$  is a total length (weight) of the defined routes;

$T_s$  is a total time of conducting search activities;

$\bar{v}_{agt}$  is the average speed of the agents moving along the defined routes;

$T_{dir}$  is an established (directive) time for performing search activities;

$K_{MAS}$  is the required number of MAS agents;

$P_{dso}$  is a probability of detecting search objects according to a certain route plan of the MAS;

$\xrightarrow{T}$  is a transitive closure between the selected pair of vertices of the model graph  $\vec{G}$ .

## INTRODUCTION

In the event of a large-scale technological disaster of a certain origin in any region of the country, which could be accompanied by the release of various types of poisonous or ionizing substances, a task of searching the missing objects and/or monitoring the operational (chemical, radiation) situation in a technological disaster zone (TDZ) may arise. In this case, the bodies for obtaining data on operational situation in such a zone can be both chemical or radiation reconnaissance groups, staffed with trained personnel and appropriate special equipment, protective gear and devices, and the groups of mobile robots equipped with appropriate devices for monitoring, measuring, and recording the values of individual parameters, as well as mixed groups.

The set of such bodies constitutes a multi-agent system (MAS), which in order to perform practical missions (achieving the goals) must have a common logic of behavior, as well as the logic of conduct of its individual agents [1, 2]. It should be noted here that the struggle to preserve the life and health of rescue and search units' personnel has recently outlined a global trend towards the use of exclusively robotic MAS in dangerous zones [3, 4]. A human in such a system endows it with its intelligence, performs control, logistical support of the system and is a consumer of the results of its functioning. The materials of this article are aimed at ensuring the functioning of robotic MAS.

Endowing the MAS with a certain behavior (intelligence) in matters of rational or optimal movement of its

elements through the network structure while performing search (or monitoring) task, requires the development of certain routing methods.

In general, the problem of routing in any network object is to find some extreme paths (shortest or k-shortest, Hamiltonian) in its structure [5, 6, 7], less often trees [8] or minimal covering trees [9], etc. In this matter, the nature of the applied problem to be solved is decisive.

While search activities using a robotic MAS, the application of the above approaches to the routing of its agents is unacceptable, since:

1) Various conditions in a certain way determine the available points of input of MAS elements in the search and the number of such points, which in turn determines the general direction of the search within TDZ.

2) In order to achieve the search goals, it is necessary to check as many areas as possible (in the best case, all areas, although it is practically impossible to fulfill this requirement in one "pass" of the MAS in dense structures).

3) It is necessary that the same areas of the terrain are not revisited by different agents (implementation of the "checked and crossed out from the list" principle), which will increase the chances of the positive result of a search operation.

4) It is desirable to withstand the general direction of advancement of the entire MAS within the network object (a technological disaster zone).

**The object of the study** is to determine the set of optimal movement routes for the elements of a multi-agent system within the network object.

**The subject of the study** is the method of optimal routing of a group (swarm) of mobile robots within the structure of a fixed network for conducting search activities in a limited area.

**The purpose of the study** is to develop a method for determining the sufficient population of a multi-agent system and the optimal movement routes of its individual agents in a stationary network for the most complete examination of a technological disaster zone (any given zone based on a certain transport network).

## 1 PROBLEM STATEMENT

Routing problems are, for the most part, formalized and solved using graph theory models and methods [10]. That is why we will model the transport network in TDZ with some weighted undirected graph  $G = (P, E)$ , where  $P = \{p_1, \dots, p_v\}$  – the set containing the vertices of the graph;  $E = \{e_1, \dots, e_m\}$  – the set containing the edges of the graph;  $v$  – the number of graph vertices;  $m$  – the number of edges of the graph. At the same time, the set  $P$  is simulating the intersections of the transport network (points at which the agent decides about the direction of its further movement), and the set  $E$  – certain sections of paths connecting different intersections in the transport network of the TDZ. Each edge from the set  $E$  will have a certain weight coefficient  $w(e_{ij})$ , which will quantita-

tively characterize the edge  $e_{ij}$ . It could be a certain number of objects to be inspected located on this site; the length of the route measured in certain units; the amount of work to be done on the site, etc.

Considering the peculiarities of carrying out search activities in a certain area, the formulation of the task of determining traffic routes for the MAS will look as follows.

Assume that the undirected weighted graph  $G=(P,E)$  has selected subset of vertices  $A = \{a_x\}_{x=1,k}$  (vertices-sources) and a subset of vertices  $B = \{b_y\}_{y=1,z}$  (vertices-sinks), where  $k$  and  $z$  are the numbers of such vertices, respectively, and  $B \notin \emptyset$ ,  $|A| \ll |P|$ ,  $|B| \ll |P|$ ,  $A \cap B = \emptyset$ . Weighting coefficients  $w(e_{ij})$  on the edges of the graph  $G=(P,E)$  quantitatively characterize the corresponding area of the TDZ.

It is necessary to find all possible edge-simple longest paths  $M_{i=1,n} = (a_x, p_1, p_2, \dots, p_q, b_y)$ , connecting vertices from subsets  $A$  and  $B$ , and the totality of which satisfies the following target function:

$$F = \sum_{i=1}^n L_i \rightarrow \max, \quad (1)$$

under the conditions:

$$T_s \leq T_{dir}, \quad (2)$$

$$n = f[G(P,E), |A|, |B|, Direct] \quad (3)$$

$$e_{ij} \in M_1 \cap e_{ij} \in M_2 \cap \dots \cap e_{ij} \in M_n = \emptyset, \quad (4)$$

$$k_{wvs} = 1. \quad (5)$$

Therefore, the solution of the problem (1)–(5) requires searching within the structure of the initial weighted graph  $G=(P,E)$  of all extremal (edge-simple longest) paths between defined subsets of its vertices.

An edge-simple path is a path in which each edge (arc) of the graph occurs only once [10].

## 2 REVIEW OF THE LITERATURE

At present, the theoretical basis for the functioning of robotic vehicles is the theory of algorithms. The development of the theory of algorithms takes place in two directions: first, the expansion of the range of practical problems solved by the existing algorithms; secondly, the development and improvement of algorithms for solving new problems that arise during the creation and functioning of the MAS. Well-developed graph theory plays a leading role in routing of individual MAS agents.

A well-known and studied problem of graph theory having numerous practical applications is the problem of Hamiltonian paths, that is, whether there is a simple path in a graph in which each vertex of the graph occurs exactly once. Such a graph is called Hamiltonian. In the case

when the graph does not contain a Hamiltonian path, in some applications it makes sense to search for a path of a maximum length in the graph. Finding such a path is known as the longest path problem. Like finding the Hamiltonian path, finding the longest path is also a difficult task.

The longest path problem is NP-complete on every class of graphs in which the Hamiltonian path problem is also NP-complete. Thus, in [11] the authors prove that even if a graph has a Hamiltonian path, the problem of finding a path of the length  $n - n^\epsilon$  for some  $\epsilon < 1$  is a NP-complete, where  $n$  – is the number of vertices of the initial graph. The authors claim that there is no polynomial approximation algorithm with a constant factor for the longest path problem unless  $P = NP$  [11]. Similar research results are also given in works [12, 13, 14, 15, 16].

It should be noted here that the Hamiltonian path problem is NP-complete on general graphs [17, 18] and remains NP-complete even when restricted to some small classes of graphs, such as splitting graphs [19], chordal bipartite graphs, strongly chordal graphs [20], directed path graphs [21], circular graphs [22], planar graphs [18] and grid graphs [23].

Polynomial solutions of this problem are known only for certain classes of graphs. Such algorithms were developed for graphs of intervals and presented in studies [24, 25, 26, 27], for doubly convex graphs – in a study [28], for graphs of arcs of circles – in a study [27] and graphs of comparability – in a study [29].

Unlike the Hamiltonian path problem, several polynomial complexity algorithms are known for the longest path problem that work with the structures of tree-type and some classes of graphs. A linear algorithm for finding the longest path on a tree-structure was proposed by Dijkstra in 1960, a formal description of which can be found in [30]. Later, based on the results of improving the Dijkstra algorithm for trees, the authors of [31] have solved the problem of finding the longest path for weighted trees and block graphs with a linear calculation time, and for cactuses with a polynomial calculation time –  $O(n^2)$ , where  $n$  – the number of vertices of the initial graph. Recently, polynomial algorithms were proposed that solve the problem of finding the longest path on bipartite graphs with computational complexity  $O(n)$  [32], on Ptolemaic graphs with computational complexity  $O(n^5)$  [33]. In [34], the authors presented their polynomial algorithm for interval graphs, which is based on the idea of dynamic programming and has a computational complexity of  $O(n^4)$ . In [35], the authors proposed a polynomial algorithm in which they also used the dynamic programming approach but applied lexicographic search in depth (the so-called LDFS graph routing) for co-comparable graphs. The computational complexity of such an algorithm is also limited to  $O(n^4)$ .

The analysis of the literature shows that the problem in the formal statement (1)–(5) was not posed or solved by anyone. If the approach is not strict, then the task is to

find in the structure of the model weighted graph  $G = (P, E)$  every (or  $k$ , where  $k > 0$ ) of the longest edge-simple paths between the selected subsets of the vertices of the graph, which lie on opposite edges of it and the total sum of the weights of which (paths) maximizes the target function  $F$ , expression (1). At the same time, an additional and mandatory condition for the sought route plan is the requirement that there are no common edges in different paths (routes). Our article is dedicated to solving this problem.

### 3 MATERIALS AND METHODS

It is quite clear that the solutions of the problem (1)–(5) formulated above require not only estimates of the lengths of the critical paths, but also the critical paths themselves, that is a consecutive set of edges that make them (the paths). It should be noted that such identification capabilities, in contrast to the well-known in graph theory algorithms Floyd-Warshall [36, 37, 38], Shimmel [39], Danzig [5] and some others, which solve the problem of finding the lengths of extreme paths, the Dijkstra’s algorithm provides.

In addition, to solve the problem (1)–(5), it is necessary to use the modified Dijkstra algorithm, since it (the problem) in this formulation belongs to the class of NP-complete and, therefore, it cannot be solved in polynomial time [41].

**Modification of the initial structure.** The problem of NP-completeness of the task of finding the longest paths (paths of the greatest total weight) in the structure of a network object is associated with the possible presence of cycles, which will lead to an unjustified increase in the total weight of the searched path (so-called “looping”), and therefore to the inability to adequately identify the path itself in the future. To solve such a “problem” when determining maximum (search for the longest paths), usually the initial undirected graph  $G = (P, E)$  is represented

in the form of an oriented graph  $\vec{G} = \left( \vec{P}, \vec{E} \right)$  without cy-

cles. For this, the edges of the graph are directed along the general search direction, which is determined based on the nature of the practical problem to be solved (see Fig. 1). As a result of this transformation, the edges of the graph become arcs, that is, they get directionality.

Since a group of mobile robots during search activities moves from one side of the TDZ to the opposite (which can be imagined in the form of a “wave” that passes through this zone), then such a general direction also exists.

As can be seen from Fig. 1, there are no cycles in the structure of the oriented graph, which makes it possible to apply the modified Dijkstra algorithm to it to find the longest paths (paths of the greatest total weight).

**Modification of initial data.** The next problematic point that needs to be solved is the representation of the weight coefficients of the edges when solving the problem of maximizing the lengths of the routes by which the ele-

ments of the MAS advance (search for the longest paths). This problem appears because in Dijkstra’s algorithm, at each iteration, the value of the accumulated total weight for some vertex is compared with the estimate of the total weight that this vertex can receive through another edge (arc) and the smallest of these values is chosen. The working expression of Dijkstra’s algorithm in its classical form is as follows [5]:

$$d(x) := \min \{d(x), d(y) + w(y, x)\}. \quad (6)$$

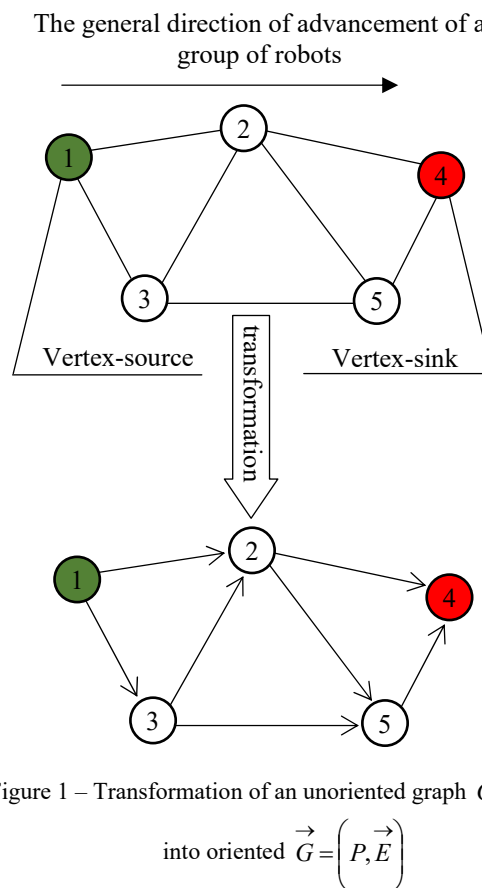


Figure 1 – Transformation of an unoriented graph  $G = (P, E)$  into oriented  $\vec{G} = \left( \vec{P}, \vec{E} \right)$

Since the problem (1)–(5) is solved for the maximum of the function, the application of the classic Dijkstra algorithm according to the expression (6) will lead to incorrect results – the search for the shortest paths. To solve such a problem, we suggest representing the weighting coefficients of the arcs of the graph  $\vec{G} = \left( \vec{P}, \vec{E} \right)$  as numbers from the negative domain and applying the following modification of the working expression of the algorithm, namely:

$$-d(x) := \min \{-d(x), -d(y) + (-w(y, x))\}. \quad (7)$$

As a result of applying expression (7), the current negative score will be compared with an alternative also negative score, and if the alternative score is lower, then the current vertex will be assigned exactly that score. In



other words, if in the structure of a weighted directed graph it is possible to increase the length of the current path due to the addition of a certain arc with a negative weight, then such an arc will be added, which will ensure that the longest path is found.

It should be noted here that estimates of path lengths will be presented in negative form, which will in no way interfere with their subsequent identification. In the future, the obtained estimates will be taken by the module, which will return them to their original physical meaning, Fig. 2.

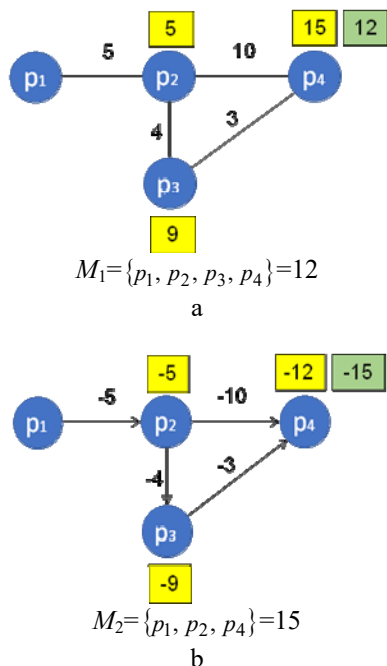


Figure 2 – The result of applying classical and modified Dijkstra methods to find the extreme paths: a – finding the shortest path  $M_1$  (classical method); b – finding the longest path  $M_2$  (modified method)

As can be seen from Fig. 2, b that between the vertices  $p_1$  and  $p_4$  there are two alternative ways:  $\{p_1, p_2, p_3, p_4\}$  and  $\{p_1, p_2, p_4\}$ . At the vertex  $p_4$  you need to decide which of these paths is considered the longest. Two arcs are incidental to it:  $(p_3, p_4)$  and  $(p_2, p_4)$ . Since the vertex  $p_3$  has a label of  $-9$  (the total weight of the path connecting this vertex to the source vertex  $p_1$ ), the vertex  $p_4$  through the arc  $(p_3, p_4)$  will receive a label of  $-12$ . Through an incident arc  $(p_2, p_4)$  having weight of  $-10$ , the vertex  $p_4$  will receive a label of  $-15$ , because the vertex  $p_2$  already has a label of  $-5$ . Therefore, the length of the path  $\{p_1, p_2, p_3, p_4\} = -12$ , and the path  $\{p_1, p_2, p_4\} = -15$ . Based on expression (7), the second of them is unambiguously chosen as the longest path. In the future, taking the received estimate of the length of the path by the module, we consider 15 units of the conditional length of the path.

On more branched structures, the process continues until all arcs are analyzed, and all vertices (including ver-

© Batsamut V. M., Hodlevskiy S. O., 2023  
 DOI 10.15588/1607-3274-2023-1-13

tices-sinks) receive the corresponding conditional weight labels according to the expression (7). The identification of the extreme paths takes place on the reverse course, just as in the classic Dijkstra method.

**Application of dynamic programming.** The characteristics of edge-simple extreme paths are of particular importance in the context of search activities in TDZ, since it is undesirable for different agents of the MAS to move along paths that have common sections (arcs), even if the number of such sections is relatively small. In this sense, the movement of different MAS agents needs to be “separated” and redirected by different ways to increase the overall effectiveness of search activities (Fig. 3). The only place where the routes of different agents can intersect is the intersection (vertices in the structure of the initial model graph), which in no way hinders their simultaneous progress along the assigned routes and, in this part, does not reduce the overall effectiveness of search activities.

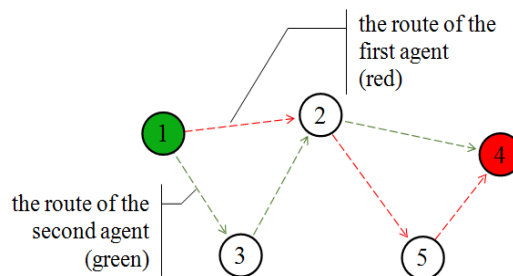


Figure 3 – Separated traffic routes of various MAS agents (option)

When solving the problem (1)–(5), to obtain a set of routes that do not have common sections (according to the additional condition (4) of the given problem) we propose to use the method of dynamic programming, which is to divide some general problem  $\hat{Z}$  on a number of subproblems  $Z', Z'', \dots, Z^n$ , where  $n$  – the number of such subproblems, and to find such their corresponding solutions  $R', R'', \dots, R^n$ , that  $R' \cup R'' \cup \dots \cup R^n$  will be the solution to a general problem  $\hat{Z}$ . At the same time, it is considered if separate solutions  $R', R'', \dots, R^n$ , are rational (or optimal), then the solutions to the general problem  $\hat{Z}$  are also rational (or optimal).

Using dynamic programming approaches and solving the problem of finding all edge-simple longest paths, we propose to split the structure of the initial graph  $\vec{G}$  into a sequence of substructures  $\vec{G}', \vec{G}'', \dots, \vec{G}^n$  and search for the appropriate junctions (routes)  $M_1, M_2, \dots, M_n$  on each of them. The set of such routes will make up the general route plan for the MAS in a technological disaster zone.

Splitting the initial graph  $\vec{G}$  into a set of substructures  $\vec{G}', \vec{G}'', \dots, \vec{G}^n$  will be based on the iterative removal from the current substructure of the arcs included in the composition of the extreme path found on this substructure.

The process of splitting the initial graph  $\tilde{G}$  will continue until there remains no path (transitive closure [42]) connecting the vertices from the sets  $A$  and  $B$  (see the statement of the problem), or until the fulfillment of an-

other (additional) condition at the choice of the decision maker.

The developed method is structurally presented in Fig. 4.

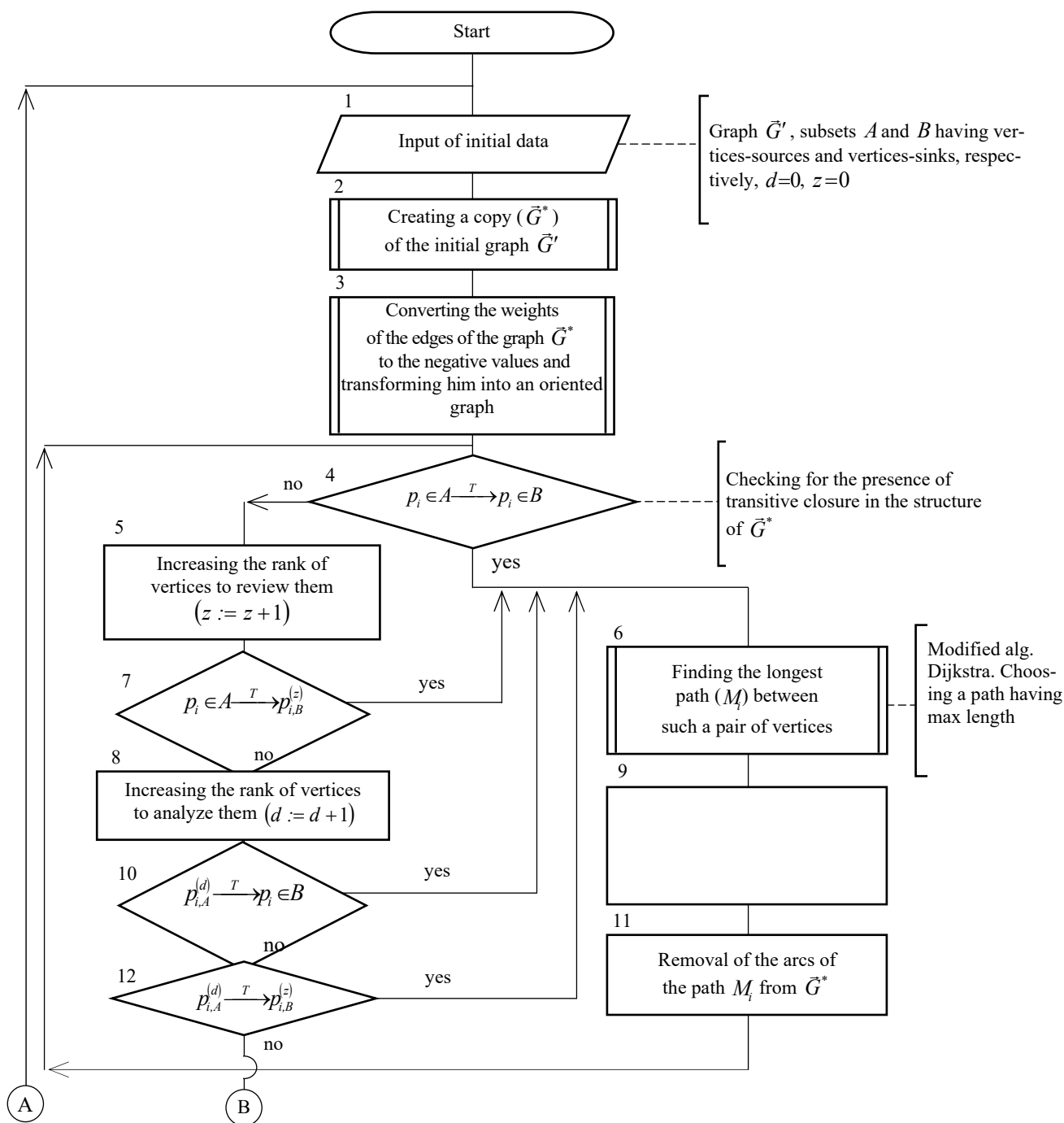


Figure 4 – Block diagram of the method of routing a group of mobile robots on a fixed network to search the missing objects in a technological disaster zone

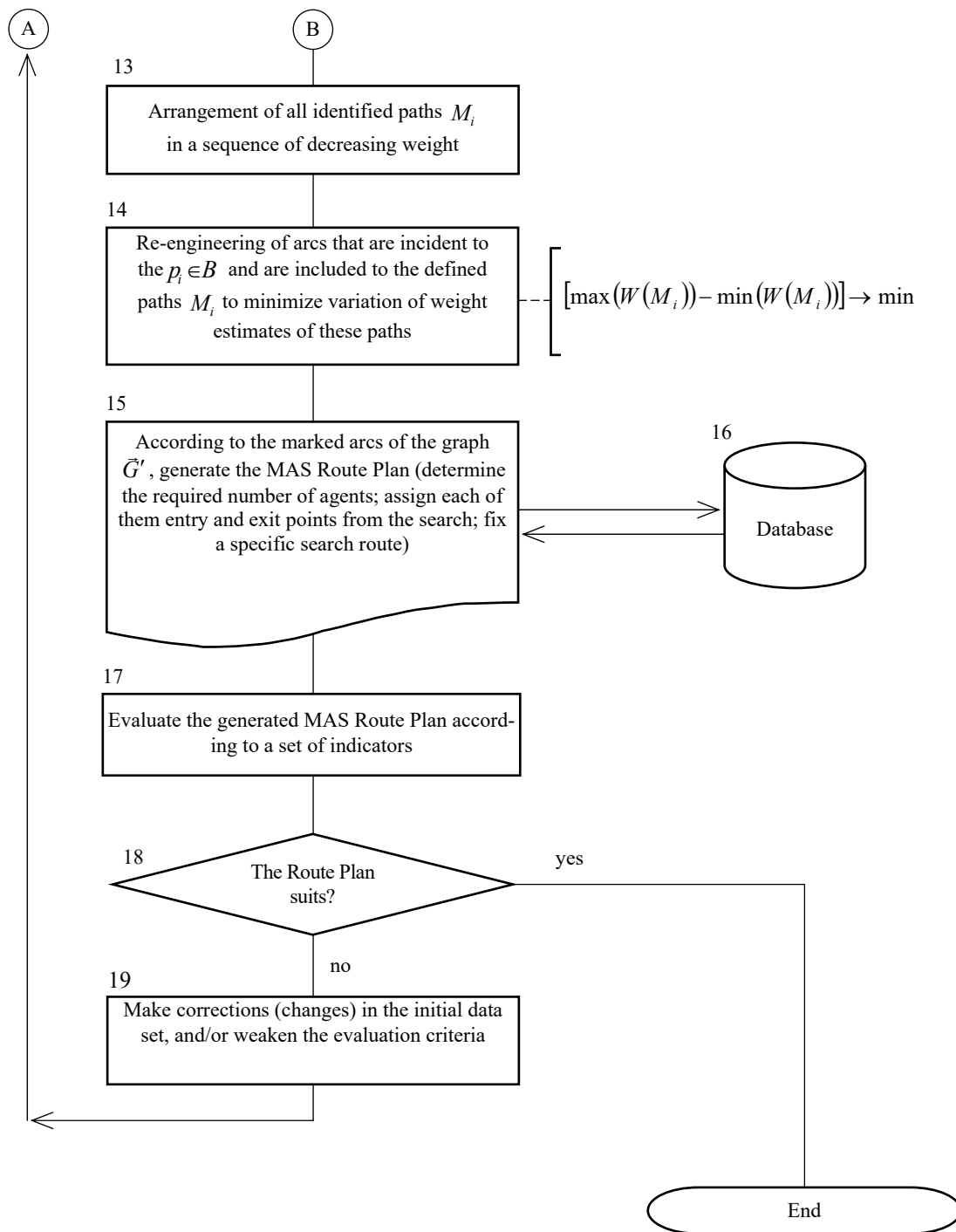


Figure 4, sheet 2

#### 4 EXPERIMENTS

Let's assume that the TDZ transport network is modeled as weighted  $G'$  (Fig. 5).

Let's give the edges of the graph a general direction from the vertex  $p_5$  towards the opposite edge of the graph. Thus, subset  $A$  will contain only one vertex  $p_5$ . We assume the subset  $B$  consists of vertices  $p_3$  and  $p_{10}$ .

Let's represent the weights of the edges as negative values.

Step I. Let's find the longest path from the vertex  $p_5$  to any vertex from the subset  $B$ . The current weight estimates obtained due to the modified Dijkstra method and the identified paths themselves are presented in Fig. 6.

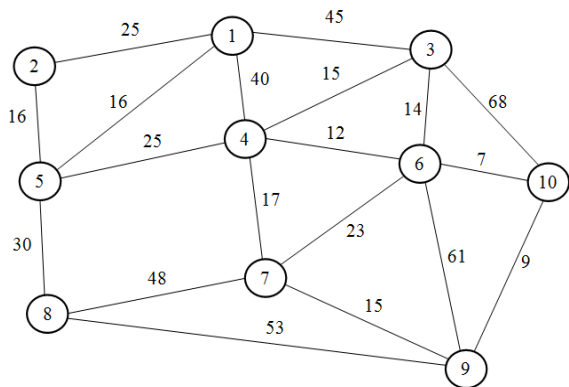


Figure 5 – Weighted graph  $\vec{G}'$  simulating the transport network in a technological disaster zone (the weights of edges are conditional lengths of the corresponding communications)

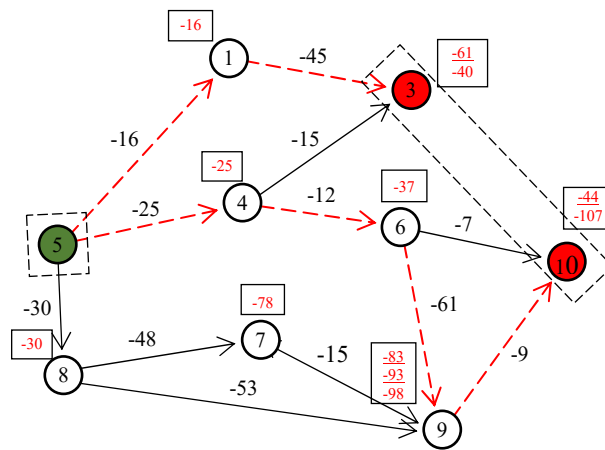


Figure 7 – Weighted graph  $\vec{G}''$

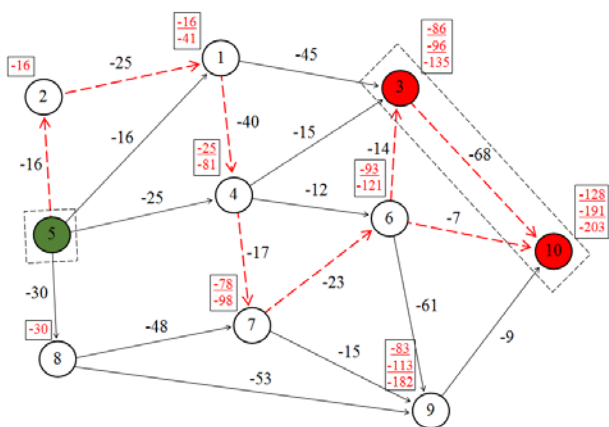


Figure 6 – Initial weighted directed graph  $\vec{G}'$

We obtained two paths (routes) to the vertices from the subset  $B$ , namely:  $M_1 = \{p_5, p_2, p_1, p_4, p_7, p_6, p_3, p_{10}\}$  and  $M_2 = \{p_5, p_2, p_1, p_4, p_7, p_6, p_{10}\}$ . The absolute weight of route  $M_1$  is greater than that of route  $M_2$ , because  $(|-203| > |-135|)$ . Therefore, the route  $M_1$  is included in the MAS route plan.

Step II. Subsequently, all arcs that make up the route  $M_2$  are removed from the structure  $\vec{G}'$ . Along with them, all hanging vertices arising from the removal of arcs are removed. As a result, we get the following structure  $\vec{G}''$  on which the next longest paths are to be searched (see Fig. 7).

After applying the modified Dijkstra method to the structure of  $\vec{G}''$ , two paths (routes) to the vertices from the subset  $B$  are obtained, namely:  $M_3 = \{p_5, p_1, p_3\}$  and  $M_4 = \{p_5, p_4, p_6, p_9, p_{10}\}$ . The absolute weight of route  $M_4$  is greater than that of route  $M_3$ , because  $(|-107| > |-61|)$ . Therefore, route  $M_4$  is included in the MAS route plan.

Step III. All arcs that make up the route  $M_4$  are to be removed from structure  $\vec{G}''$ . All hanging vertices are also to be removed. We obtain the following structure  $\vec{G}'''$ , on which the longest path to any vertex from subset  $B$  is also being searched (see Fig. 8). Such a route is the only route  $M_5 = \{p_5, p_1, p_3\}$  and its weight is  $|-61|$ . Route  $M_5$  is to be included into the MAS route plan.

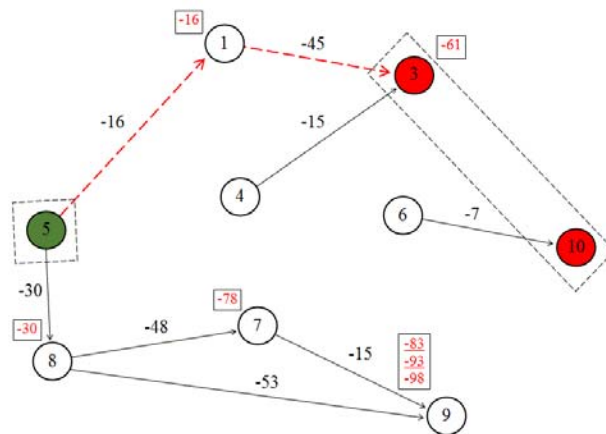


Figure 8 – Weighted graph  $\vec{G}'''$

Step IV. All arcs that make up the route  $M_5$  from the structure of  $\vec{G}'''$  are to be deleted. Hanging vertices are also removed. As a result, we get the following structure of  $\vec{G}''''$  (see Fig. 9). We can see from the figure that in this structure there are no paths connecting vertex  $p_5$  to any vertex from subset  $B$ . So, based on the results of this step, the problem (1)–(5) has been basically solved. On the initial structure, all edge-simple longest paths between the defined sets of vertices in the structure of the initial network object  $\vec{G}'$  have been determined. We will name such paths as first-level paths because they directly connect vertices from subsets  $A$  and  $B$ .

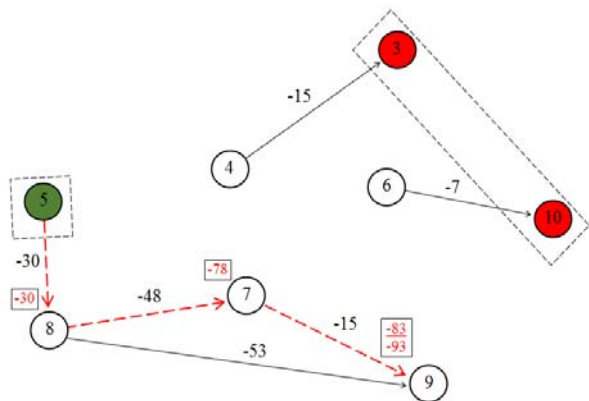


Figure 9 – Weighted graph  $\vec{G}''''$

However, vertex  $p_5$  has another unused incident arc  $(p_5, p_8)$ . Let's find the maximum path to any vertex (in this case, to one that does not belong to subset  $B$ ). We will consider such paths as paths of the second level: they connect the initial vertex-source with some vertex that is adjacent to some vertex from the subset  $B$ . Such paths can also be included into the MAS route plan, thereby increasing its efficiency.

In structure  $\vec{G}''''$ , such a route is the only route  $M_6 = \{p_5, p_8, p_7, p_9\}$  with total weight  $|-93|$ . Route  $M_6$  is also included into the MAS route plan.

So, as a result of successive splitting of the initial graph  $\vec{G}'$ , namely:  $\vec{G}' \rightarrow \vec{G}'' \rightarrow \vec{G}''' \rightarrow \vec{G}''''$ , four edge-simple longest paths (those that do not have mutual arcs) were found in its structure, namely:  $M_1 = \{p_5, p_2, p_1, p_4, p_7, p_6, p_3, p_{10}\}$ ;  $M_4 = \{p_5, p_4, p_6, p_9, p_{10}\}$ ;  $M_5 = \{p_5, p_1, p_3\}$ ;  $M_6 = \{p_5, p_8, p_7, p_9\}$ . The paths and their absolute weights are presented in the figure (Fig. 10).

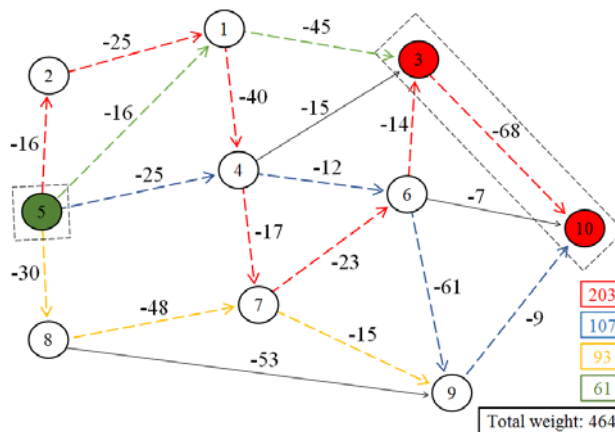


Figure 10 – The set of searched edge-simple longest paths in the structure of the initial network object  $\vec{G}'$  (marked with different colors)

### 5 RESULTS

The scheme of routes presented in Fig. 10, contains the following information:

- to carry out search activities within the TDZ, whose transport network is modeled by graph  $\vec{G}'$ , it is sufficient to have four agents as part of the MAS.
- for each MAS agent, its starting point (boundary) of entering the search and the final point (boundary) of exiting the search are defined.
- each MAS agent within the TDZ is assigned a specific search (movement) route.
- routes of movement of various MAS agents do not cross sections (arcs).
- in total, the traffic routes of the MAS are optimized both in terms of the direction of traffic and the length of the routes.

The assessments of the developed MAS route plan, based on the totality of the proposed indicators, are presented in Table 1.

Table 1 – Assessments of the developed MAS route plan according to a set of indicators

No	Indicator (parameter)	Indicator value
1	Required number of MAS agents, $K_{MAS}$ , (units):  $K_{MAS} \cong n$ ,  where $n$ – the number of defined (selected) search routes	4
2	MAS route plan, $PL_{MAS}$ :  $PL_{MAS} = \{M_1, M_2, \dots, M_n\}$ ,  where $n$ – the number of defined (selected) search routes	$M_1 = \{p_5, p_2, p_1, p_4, p_7, p_6, p_3, p_{10}\}$ ; $M_4 = \{p_5, p_4, p_6, p_9, p_{10}\}$ ; $M_5 = \{p_5, p_1, p_3\}$ ; $M_6 = \{p_5, p_8, p_7, p_9\}$

Table 1 – Assessments of the developed MAS route plan according to a set of indicators (continuation)

№	Indicator (parameter)	Indicator value
3	Route length, $L_i$ , (km)	$L_1 = 203; L_4 = 107; L_5 = 61; L_6 = 93$
4	Total length (weight) of routes, $L_{PL}$ , (km): $L_{PL} = \sum_{i=1}^n L_i,$ where $n$ – the number of defined (selected) search routes	464
5	Assessment of the search time in TDZ, $T_s$ , (hrs.): $T_s = \frac{\max \{L_i\}}{\bar{v}_{agt}}, i = 1 \dots n,$ where $\bar{v}_{agt}$ – the average speed of the agents moving along the defined routes, km/h.	If: $\bar{v}_{agt} = 5$ km/h. $T_s \approx 40,6$ hours
6	Probability of finding search objects in case of implementation of the MAS route plan, $P_{dso}$ : $P_{dso} = \frac{L_{PL}}{L_{total}},$ where $L_{total}$ – the total length of the transport network in the TDZ	If: $L_{total} = 539$ km $P_{dso} = 0,86$
7	Probability of finding search objects within the prescribed time for conducting search activities, $P_{dso}(T_s \leq T_{dir})$ : $P_{dso}(T_s \leq T_{dir}) = 1 - e^{-\left(\frac{T_{dir}}{T_s}\right)},$ where $T_{dir}$ – the prescribed time	If: $T_{dir} = 30$ hours $P_{dso}(T_s \leq T_{dir}) = 0,52$

## 6 DISCUSSION

The set of approaches proposed in the article, together with the modified Dijkstra’s method, allowed to develop a method that searches in the structure of any network object, between subsets of its vertices, all existing edge-simple longest paths, which allows to organize the routing of MAS agents in a certain area by paths that do not intersect their individual sections.

At the same time, it is worth noting that the structure of the paths themselves, as well as their number, will depend on the defined (prescribed) points of entry into the search and points of withdrawal from the search, which allows, when planning search activities, to quickly consider several options for future actions and choose the most appropriate one (to order the options by the degree of their advantage).

In addition, having carefully studied the scheme of routes presented in Figure 10, it is possible to conclude that it is expedient to reengineer the routes to reduce the variation of values by their absolute weight (see Fig. 4, Block 14). The idea of reengineering is as follows. If there is an arc (edge) connecting them between any pair of vertices from a subset  $B$ , and this arc (edge) is included in some extreme path, and this path has the maximum (higher) weight score among other paths, then this arc (edge) should be included in the path with the minimum (lower) weight. It should be noted that based on the results of using the proposed method, the minimally required number of MAS agents to conduct search activities

in an area with a certain transport network could be found.

It is obvious that in the real structures, and hence in the structures which are denser and having more vertices, the number of iterations to split the initial structure and, accordingly, to find the edge-simple longest path, can be close to the very dimension of the initial structure itself. Therefore, the computational complexity of the combinatorial algorithm that implements the developed method will be determined by the computational complexity of its “basic element” – the Dijkstra’s algorithm, which is estimated as  $O(n^2)$  [40], as well as the computational complexity of the algorithm that determines the presence of transitive closure between vertices from subsets  $A$  and  $B$ , which is evaluated as  $O(n^3)$  [41]. Considering the strengths of subsets  $A$  and  $B$ , the total computational complexity of the combinatorial algorithm could be estimated as  $O(|A| \cdot |B| \cdot (n^3 + n^2))$ , where  $|A|$  and  $|B|$  – the strengths of the corresponding subsets;  $n$  – the number of vertices of the model graph  $G$ .

The obtained polynomial estimate of computational complexity is quite acceptable for the use of such an algorithm in a real-time framework.

## CONCLUSIONS

The article solves the actual scientific and applied problem of finding all possible edge-simple longest paths

between defined subsets of vertices of an initial undirected graph.

**The scientific novelty** of the developed method is as follows:

1) the representation of the weights of the edges of the initial model graph by the negative values, which allows finding the longest paths between a specified pair of vertices using the classical Dijkstra method;

2) the application of the dynamic programming method, which makes it possible to find the longest paths  $M_1, M_2, \dots, M_n$  in the set of obtained substructures  $\bar{G}', \bar{G}'', \dots, \bar{G}^n$  of the initial model graph  $\bar{G}$ , which will constitute the complete combination of the edge-simple longest paths.

Since the basic element of the developed method is the Dijkstra's method, which belongs to the accurate class, it can be assumed that the developed method is also accurate.

**The practical value** of the method is that its application significantly simplifies the process of developing effective route plans for the elements of the MAS in a particular transport network. The projection of general theoretical and methodological statements and conclusions made during the study on the problem of routing of the MAS elements in a certain area, allows to make the search activities effective and quickly develop several options for search actions. The developed method will have a practical use if it is implemented based on a geoinformation system.

**A promising direction for further research** is the development of routing methods of the MAS performing search activities in a non-stationary network, as well as the development of a set of indicators and criteria for prompt decision-making regarding the optimal (rational) bypassing by agents of the obstacles that suddenly appear on the route path.

#### ACKNOWLEDGEMENTS

This article highlights one of the results obtained by the authors in 2020–2021 while implementing the research project (state registration number 0120U002173) at the Research Center of the National Academy of the National Guard of Ukraine. The authors are grateful to their colleagues for their support during the research and active participation in the discussion of the results. All authors declare that they have neither financial support nor obligations.

#### REFERENCES

1. Konovalenko O., Brusentsev V. Multi-agent management and decision support systems, *Bulletin of the National Technical University "KhPI". Ser.: Mechanical engineering and CAD*, 2019, Vol. 1, pp. 18–27. DOI: 10.20998/2079-0775.2019.1.03
2. Kravari K., Bassiliades N. A Survey of Agent Platforms, *Journal of Artificial Societies and Social Simulation*, 2015, Vol. 18, № 1, pp. 1–18. DOI: 10.18564/jasss.2661
3. Schurr N., Schurr N., Marecki J. et al. The Future of Disaster Response: Humans Working with Multiagent Teams using DEFACTO, *Conference: AI Technologies for*

- Homeland Security : Papers from the 2005 AAAI Spring Symposium, Technical Report SS-05-01*. Stanford, California, USA, March 21–23, 2005 : proceedings, Menlo Park, California : The AAAI Press, 2005, pp. 9–16.
4. Ayanian N. DART: Diversity-enhanced Autonomy in Robot Teams, *The International Journal of Robotics Research*. – 2018, pp. 1–8. DOI: 10.1177/0278364919839137.
5. Minieka E. Optimization algorithms on networks and graphs. Moscow, Mir, 1981, 323 p.
6. Held M., Karp R. The Travelling-Salesman Problem and Minimum Spanning Trees: Part II, *Math. Programming*, 1971, Vol. 1, pp. 6–25. DOI:10.1007/BF01584070
7. Bellman R. On a Routing Problem, *Quarterly of Applied Mathematics*, 1958, Vol. 16 (1), pp. 87–90. DOI: 10.1090/qam/102435
8. Hadjiconstantinou E., Christofides N., Mingozi A. A new exact algorithm for the vehicle routing problem based on q-paths and k-shortest paths relaxations, *Annals of Operations Research*, 1995, Vol. 61, pp. 21–43. DOI:10.1007/BF02098280
9. Christofides N., Mingozi A., Toth P. Exact algorithm for the Vehicle Routing Problem Based on Spanning Tree and Shortest Paths Relaxations, *Mathematical Programming*, 1981, Vol. 20, No. 1, pp. 255–282. DOI:10.1007/BF01589353
10. Christofides N. Theory of graphs. Algorithmic approach. Moscow, Mir, 1978, 432 p.
11. Karger D., Motwani R., Ramkumar G.D.S. On approximating the longest path in a graph, *Algorithmica*, 1997, Vol. 18, pp. 82–98. DOI: 10.1007/BF02523689
12. Feder T., Motwani R. Finding large cycles in Hamiltonian graphs, *16th annual ACM-SIAM Symp. on Discrete Algorithms (SODA), Vancouver, 23–25 January 2005 : proceedings*. Philadelphia, United States, Society for Industrial and Applied Mathematics, 2005, pp. 166–175. DOI: 10.5555/1070432
13. Gabow H. N. Finding paths and cycles of super polylogarithmic length, *36th annual ACM Symp. on Theory of Computing (STOC), Chicago, 13–16 June 2004 : proceedings*. New York, United States : Association for Computing Machinery, 2004, pp. 407–416. DOI: 10.1145/1007352.1007418
14. Gabow H. N., Nie S. Finding long paths, cycles, and circuits, *19th annual International Symp. on Algorithms and Computation (ISAAC), Gold Coast, Australia, December 2008 : proceedings*. Berlin, Springer-Verlag, 2008, pp. 752–763. DOI: 10.1007/978-3-540-92182-0\_66
15. Vishwanathan S. An approximation algorithm for finding a long path in Hamiltonian graphs, *11th annual ACM-SIAM Symp. on Discrete Algorithms (SODA), San Francisco, 9–11 January 2000 : proceedings*. Philadelphia, United States : Society for Industrial and Applied Mathematics, 2000, pp. 680–685. DOI: 10.5555/338219
16. Zhang Z., Li H. Algorithms for long paths in graphs, *Theoretical Computer Science*, 2007. Vol. 377, Issue 1–3, pp. 25–34. DOI: 10.1016/j.tcs.2007.02.012
17. Garey M. R., Johnson D. S. Computers and Intractability: A Guide to the Theory of NP-completeness. New York, W.H. Freeman, 1979, 340 p.
18. Garey M. R., Johnson D. S., Tarjan R. E. The planar Hamiltonian circuit problem is NP-complete, *SIAM J. Computing*, 1976. Vol. 5, pp. 704–714. DOI: 10.1137/0205049
19. Golumbic M. C. Algorithmic Graph Theory and Perfect Graphs, Vol. 57 : *Annals of Discrete Mathematics*. 2nd

- Edition*. Amsterdam, North-Holland Publishing Co., 2004, 592 p. ISBN: 9780080526966
20. Müller H. Hamiltonian circuits in chordal bipartite graphs, *Discrete Math.*, 1996, Vol. 156, pp. 291–298. DOI: 10.1016/0012-365X(95)00057-4
  21. Narasimhan G. A note on the Hamiltonian circuit problem on directed path graphs, *Information Processing Letters*, 1989, Vol. 32, pp. 167–170. DOI: 10.1016/0020-0190(89)90038-0
  22. Damaschke P. The Hamiltonian circuit problem for circle graphs is NP-complete, *Information Processing Letters*. – 1989, Vol. 32, pp. 1–2. DOI: 10.1016/0020-0190(89)90059-8
  23. Itai A., Papadimitriou C. H., Szwarcfiter J. L. Hamiltonian paths in grid graphs, *SIAM J. Computing*, 1982, Vol. 11, pp. 676–686. DOI: 10.1137/0211056
  24. Arikati S. R., Pandu Rangan C. Linear algorithm for optimal path cover problem on interval graphs, *Information Processing Letters*, 1990, Vol. 35, pp. 149–153. DOI: 10.1016/0020-0190(90)90064-5
  25. Bertossi A. A. Finding Hamiltonian circuits in proper interval graphs, *Information Processing Letters*, 1983, Vol. 17, pp. 97–101. DOI: 10.1016/0020-0190(83)90078-9
  26. Chang M. S., Peng S. L., Liaw J. L. Deferred-query: An efficient approach for some problems on interval graphs, *Networks*, 1999, Vol. 34, Issue 1, pp. 1–10. DOI: 10.1002/(sici)1097-0037(199908)34:1<1::aid-net1>3.0.co
  27. Damaschke P. Paths in interval graphs and circular arc graphs, *Discrete Math.*, 1993, Vol. 112, pp. 49–64. DOI: 10.1016/0012-365X(93)90223-G
  28. Asdre K., Nikolopoulos S. D. The 1-fixed-endpoint path cover problem is polynomial on interval graphs, *Algorithmica*, 2009, Vol. 58, pp. 679–710. DOI: 10.1007/s00453-009-9292-5
  29. Damaschke P., Deogun J. S., Kratsch D. et al. Finding Hamiltonian paths in cocomparability graphs using the bump number algorithm, *Order*, 1992, Vol. 8, pp. 383–391. DOI: 10.1007/bf00571188
  30. Bulterman R., Sommen F. van der, Zwaan G. et al. On computing a longest path in a tee, *Information Processing Letters*, 2002, Vol. 81, pp. 93–96. DOI: 10.1016/S0020-0190(01)00198-3
  31. Uehara R., Uno Y. Efficient algorithms for the longest path problem, *15th annual International Symp. on Algorithms and Computation (ISAAC)*, Hong Kong, December 2004 : *proceedings*. Berlin, Springer-Verlag, 2004, pp. 871–883. DOI:10.1007/978-3-540-30551-4\_74
  32. Uehara R., Valiente G. Linear structure of bipartite permutation graphs and the longest path problem, *Information Processing Letters*, 2007, Vol. 103, pp. 71–77. DOI: 10.1016/j.ipl.2007.02.010
  33. Takahara Y., Teramoto S., Uehara R. Longest path problems on ptolemaic graphs, *IEICE Trans. Inf. and Syst.*, 2008, Vol. 91-D, pp. 170–177. DOI: 10.1093/ietisy/e91-d.2.170
  34. Ioannidou K., Mertzios G., Nikolopoulos S. The Longest Path Problem has a Polynomial Solution on Interval Graphs, *Algorithmica*, 2011, Vol. 61, pp. 320–341. DOI: 10.1007/s00453-010-9411-3
  35. Mertzios G. B., Corneil D. G. A Simple Polynomial Algorithm for the Longest Path Problem on Cocomparability Graphs, *SIAM Journal on Discrete Mathematics*, 2012, Vol. 26, Issue 3, pp. 940–963. DOI: 10.1137/100793529
  36. Floyd R. Algorithm 97: Shortest Path, *Communications of the ACM*, 1961, Vol. 5 (6), pp. 344–348. DOI: 10.1145/367766.368168
  37. Hougardy S. The Floyd-Warshall algorithm on graphs with negative cycles, *Information Processing Letters*, 2010, Vol. 110 (8–9), pp. 279–281. DOI: 10.1016/j.ipe.2010.02.001
  38. Warshall S. Algorithm on Boolean matrices, *Journal of the ACM*, 1962, Vol. 9 (1), pp. 11–12. DOI: 10.1145/321105.321107
  39. Shimbel A. Structural parameters of communication networks, *Bulletin of Mathematical Biophysics*, 1953, Vol. 15 (4), pp. 501–507. DOI: 10.1007/BF02476438
  40. Dijkstra E. W. A note on two problems in connexion with graphs, *Numerische Mathematik*, 1959, Vol. 1, Issue 1, pp. 269–271. DOI:10.1007/BF01386390
  41. Lipsky V. Combinatory for programmers, Moscow, Mir Press., 1988, 213 p.
  42. Batsamut V., Manzura S., Kosiak O. et al. Fast Algorithm for Calculating Transitive Closures of Binary Relations in the Structure of a Network Object, *International Journal of Computing*, 2021, Vol. 20(4), pp. 560–566. DOI: 10.47839/ijc.20.4.2444

Received 22.11.2022.  
Accepted 15.12.2022.

УДК 519.853: 658.52

## МЕТОД МАРШРУТИЗАЦІЇ ГРУПИ МОБІЛЬНИХ РОБОТІВ НА СТАЦІОНАРНІЙ МЕРЕЖІ ДЛЯ ВИКОНАННЯ ЗАВДАНЬ ПОШУКУ ЗНИКЛИХ ОБ'ЄКТІВ В ЗОНІ ТЕХНОГЕННОЇ АВАРІЇ

**Бацамут В. М.** – д-р військ. наук, професор, заступник начальника науково-дослідного центру службово-бойової діяльності Національної гвардії України Національної академії Національної гвардії України, Харків, Україна.

**Годлевський С. О.** – науковий співробітник науково-дослідного центру службово-бойової діяльності Національної гвардії України Національної академії Національної гвардії України, Харків, Україна.

### АНОТАЦІЯ

**Актуальність.** Актуальність статті обумовлюється потребою у подальшому розвитку моделей колективної поведінки систем із мультиагентною побудовою структури, у наділленні таких систем інтелектом, який забезпечує синхронізацію спільних зусиль різних агентів у ході досягнення поставлених перед системою цілей. Запропонований у статті метод усуває проблему конкуренції між різними агентами мультиагентної системи, що є важливим у ході виконання пошукових, рятувальних, моніторингових завдань у кризових районах різного характеру походження.

**Мета роботи** полягає у розробленні методу визначення достатньої чисельності мультиагентної системи та оптимальних маршрутів руху її окремих елементів на стаціонарній мережі для максимально повного обстеження зони техногенної аварії (будь-якої заданої зони, в основі якої лежить певна транспортна мережа).



**Метод.** Застосовано ідею динамічного програмування для пошуку в структурі модельного зваженого орієнтованого графа всіх можливих реберно-простих найдовших шляхів, що з'єднують директивно визначені підмножини вершин-істоків та вершин-стоків. З цієї метою застосовано модифікований метод Дейкстри. Модифікація полягає у представленні ваг дуг моделюючого орієнтованого графа значеннями з від'ємної області з подальшою роботою метода Дейкстри з цими значеннями. Після відшукування чергового реберно-простого найдовшого шляху, дуги, що його складають, фіксуються у пам'яті обчислювальної системи (у маршрутному плані) та видаляються зі структури графа і процес ітераційно повторюється. Пошук шляхів відбувається доти, поки зберігається транзитивне замкнення між вершинами, що входять до складу визначених підмножин вершин-істоків та вершин-стоків. Розроблений метод дозволяє знайти таку сукупність маршрутів руху для елементів мультиагентної системи, яка максимізує обстежену ними площу в зоні техногенної аварії (або кількість перевірених об'єктів на маршрутах руху) за одну "хвилю" пошуку, та розподіляє елементи мультиагентної системи маршрутами, що не мають спільних ділянок. Похідною застосування розробленого методу є визначення достатньої чисельності мультиагентної системи для ефективного проведення пошукових заходів у межах визначеної зони.

**Результати.** 1) Розроблено метод маршрутизації групи мобільних роботів на стаціонарній мережі для виконання завдань пошуку зниклих об'єктів в зоні техногенної аварії; 2) Формалізовано робочий вираз методу Дейкстри для пошуку в структурі мережевого об'єкту (в структурі модельного графа) шляхів найбільшої довжини; 3) Запропонована сукупність показників для комплексного оцінювання маршрутних планів мультиагентної системи; 4) Виконано верифікацію методу на тестових задачах.

**Висновки.** Проведені теоретичні дослідження та низка експериментів підтверджують працездатність розробленого методу. Рішення, що виробляються із використанням розробленого методу, є точними, що дозволяє рекомендувати його до практичного використання при визначенні в автоматизованому режимі маршрутних планів для мультиагентних систем, а також потрібної кількості агентів в таких системах для виконання необхідного обсягу пошукових завдань у певному кризовому районі.

**КЛЮЧОВІ СЛОВА:** мультиагентна система, група мобільних роботів, маршрутизація, мережевий об'єкт, зважений неорієнтований (орієнтований) граф, екстремальні шляхи, критерій оптимізації, метод.

#### ЛІТЕРАТУРА

1. Коноваленко О. Мультиагентні системи управління та підтримки прийняття рішень / О. Коноваленко, В. Брусенцев // Bulletin of the National Technical University "KhPI". Ser.: Mechanical engineering and CAD. – 2019. – № 1. – С. 18–27. DOI: 10.20998/2079-0775.2019.1.03
2. Kravari K. A Survey of Agent Platforms / K. Kravari, N. Bassiliades // Journal of Artificial Societies and Social Simulation. – 2015. – Vol. 18, № 1. – P. 1–18. DOI: 10.18564/jasss.2661
3. The Future of Disaster Response: Humans Working with Multiagent Teams using DEFACTO / [N. Schurr, N. Schurr, J. Marecki et al.] // Conference: AI Technologies for Homeland Security : Papers from the 2005 AAAI Spring Symposium, Technical Report SS-05-01, Stanford, California, USA, March 21–23, 2005 : proceedings. – Menlo Park, California : The AAAI Press, 2005. P. 9–16.
4. Ayanian N. DART: Diversity-enhanced Autonomy in Robot Teams / N. Ayanian // The International Journal of Robotics Research. – 2018. – P. 1–8. DOI: 10.1177/0278364919839137
5. Майника Э. Алгоритмы оптимизации на сетях и графах / Э. Майника. – Москва : Мир, 1981. – 323 с.
6. Held M. The Travelling-Salesman Problem and Minimum Spanning Trees: Part II / M. Held, R. Karp // Math. Programming. – 1971. – Vol. 1 – P. 6–25. DOI:10.1007/BF01584070
7. Bellman R. On a Routing Problem / R. Bellman // Quarterly of Applied Mathematics. – 1958. – Vol. 16 (1). – P. 87–90. DOI: 10.1090/qam/102435
8. Hadjiconstantinou E. A new exact algorithm for the vehicle routing problem based on q-paths and k-shortest paths relaxations / E. Hadjiconstantinou, N. Christofides, A. Mingozzi // Annals of Operations Research. – 1995. – Vol. 61. – P. 21–43. DOI:10.1007/BF02098280
9. Christofides N. Exact algorithm for the Vehicle Routing Problem Based on Spanning Tree and Shortest Paths Relaxations / N. Christofides, A. Mingozzi, P. Toth // Mathematical Programming. – 1981. – Vol. 20, No. 1. – P. 255–282. DOI:10.1007/BF01589353
10. Кристофидес Н. Теория графов. Алгоритмический подход / Н. Кристофидес. – Москва : Мир, 1978. – 432 с.
11. Karger D. On approximating the longest path in a graph / D. Karger, R. Motwani, G.D.S. Ramkumar // Algorithmica. – 1997. – Vol. 18. – P. 82–98. DOI: 10.1007/BF02523689
12. Feder T. Finding large cycles in Hamiltonian graphs / T. Feder, R. Motwani // 16th annual ACM-SIAM Symp. on Discrete Algorithms (SODA), Vancouver, 23–25 January 2005 : proceedings. – Philadelphia, United States : Society for Industrial and Applied Mathematics, 2005. – P. 166–175. DOI: 10.5555/1070432
13. Gabow H. N. Finding paths and cycles of super polylogarithmic length / H. N. Gabow // 36th annual ACM Symp. on Theory of Computing (STOC), Chicago, 13–16 June 2004 : proceedings. – New York, United States : Association for Computing Machinery, 2004, P. 407–416. DOI: 10.1145/1007352.1007418
14. Gabow H.N. Finding long paths, cycles, and circuits / H. N. Gabow, S. Nie // 19th annual International Symp. on Algorithms and Computation (ISAAC), Gold Coast, Australia, December 2008 : proceedings. – Berlin : Springer-Verlag, 2008. – P. 752–763. DOI: 10.1007/978-3-540-92182-0\_66
15. Vishwanathan S. An approximation algorithm for finding a long path in Hamiltonian graphs / S. Vishwanathan // 11th annual ACM-SIAM Symp. on Discrete Algorithms (SODA), San Francisco, 9–11 January 2000 : proceedings. – Philadelphia, United States : Society for Industrial and Applied Mathematics, 2000. – P. 680–685. DOI: 10.5555/338219
16. Zhang Z. Algorithms for long paths in graphs / Z. Zhang, H. Li // Theoretical Computer Science. – 2007. – Vol. 377, Issue 1–3. – P. 25–34. DOI: 10.1016/j.tcs.2007.02.012
17. Garey M. R. Computers and Intractability: A Guide to the Theory of NP-completeness / M. R. Garey, D. S. Johnson. – New York : W.H. Freeman, 1979. – 340 p.
18. Garey M. R. The planar Hamiltonian circuit problem is NP-complete / M. R. Garey, D. S. Johnson, R. E. Tarjan //

- SIAM J. Computing. – 1976. – Vol. 5. – P. 704–714. DOI: 10.1137/0205049
19. Golubic M. C. Algorithmic Graph Theory and Perfect Graphs / M. C. Golubic. – Vol. 57 : Annals of Discrete Mathematics. 2nd Edition. – Amsterdam : North-Holland Publishing Co., 2004. – 592 p. ISBN: 9780080526966
20. Müller H. Hamiltonian circuits in chordal bipartite graphs / H. Müller // Discrete Math. – 1996. – Vol. 156. – P. 291–298. DOI: 10.1016/0012-365X(95)00057-4
21. Narasimhan G. A note on the Hamiltonian circuit problem on directed path graphs / G. Narasimhan // Information Processing Letters. – 1989. – Vol. 32. – P. 167–170. DOI: 10.1016/0020-0190(89)90038-0
22. Damaschke P. The Hamiltonian circuit problem for circle graphs is NP-complete / P. Damaschke // Information Processing Letters. – 1989. – Vol. 32. – P. 1–2. DOI: 10.1016/0020-0190(89)90059-8
23. Itai A. Hamiltonian paths in grid graphs / A. Itai, C. H. Papadimitriou, J. L. Szwarcfiter // SIAM J. Computing. – 1982. – Vol. 11. – P. 676–686. DOI: 10.1137/0211056
24. Arikati S. R. Linear algorithm for optimal path cover problem on interval graphs / S. R. Arikati, Rangan C. Pandu. // Information Processing Letters. – 1990. – Vol. 35. – P. 149–153. DOI: 10.1016/0020-0190(90)90064-5
25. Bertossi A. A. Finding Hamiltonian circuits in proper interval graphs / A. A. Bertossi // Information Processing Letters. – 1983. – Vol. 17. – P. 97–101. DOI: 10.1016/0020-0190(83)90078-9
26. Chang M. S. Deferred-query: An efficient approach for some problems on interval graphs / M. S. Chang, S. L. Peng, J. L. Liaw // Networks. – 1999. – Vol. 34, Issue 1. – P. 1–10. DOI: 10.1002/(sici)1097-0037(199908)34:1<1::aid-net1>3.0.co
27. Damaschke P. Paths in interval graphs and circular arc graphs / P. Damaschke // Discrete Math. – 1993. – Vol. 112. – P. 49–64. DOI: 10.1016/0012-365X(93)90223-G
28. Asdre K. The 1-fixed-endpoint path cover problem is polynomial on interval graphs / K. Asdre, S. D. Nikolopoulos // Algorithmica. – 2009. – Vol. 58. – P. 679–710. DOI: 10.1007/s00453-009-9292-5
29. Finding Hamiltonian paths in cocomparability graphs using the bump number algorithm / [P. Damaschke, J. S. Deogun, D. Kratsch et al.] // Order. – 1992. – Vol. 8. – P. 383–391. DOI: 10.1007/bf00571188
30. On computing a longest path in a tee / [R. Bulterman, F. van der Sommen, G. Zwaan et al.] // Information Processing Letters. – 2002. – Vol. 81. – P. 93–96. DOI: 10.1016/S0020-0190(01)00198-3
31. Uehara R. Efficient algorithms for the longest path problem / R. Uehara, Y. Uno // 15th annual International Symp. on Algorithms and Computation (ISAAC), Hong Kong, December 2004 : proceedings. – Berlin : Springer-Verlag, 2004 – P. 871–883. DOI:10.1007/978-3-540-30551-4\_74
32. Uehara R. Linear structure of bipartite permutation graphs and the longest path problem / R. Uehara, G. Valiente // Information Processing Letters. – 2007. – Vol. 103. – P. 71–77. DOI: 10.1016/j.ipl.2007.02.010
33. Takahara Y. Longest path problems on ptolemaic graphs / Y. Takahara, S. Teramoto, R. Uehara // IEICE Trans. Inf. and Syst. – 2008. – Vol. 91-D. – P. 170–177. DOI: 10.1093/ietisy/e91-d.2.170
34. Ioannidou K. The Longest Path Problem has a Polynomial Solution on Interval Graphs / K. Ioannidou, G. Mertzios, S. Nikolopoulos // Algorithmica. – 2011. – Vol. 61. – P. 320–341. DOI: 10.1007/s00453-010-9411-3
35. Mertzios G. B. A Simple Polynomial Algorithm for the Longest Path Problem on Cocomparability Graphs / G. B. Mertzios, D. G. Corneil // SIAM Journal on Discrete Mathematics. – 2012. – Vol. 26, Issue 3, – P. 940–963. DOI: 10.1137/100793529
36. Floyd R. Algorithm 97: Shortest Path / R. Floyd // Communications of the ACM. – 1961 – Vol. 5 (6). – P. 344–348. DOI: 10.1145/367766.368168
37. Hougardy S. The Floyd-Warshall algorithm on graphs with negative cycles / S. Hougardy // Information Processing Letters. – 2010. – Vol. 110 (8–9). – P. 279–281. DOI: 10.1016/j.ipe.2010.02.001
38. Warshall S. Algorithm on Boolean matrices / S. Warshall // Journal of the ACM. – 1962. – Vol. 9 (1). – P. 11–12. DOI: 10.1145/321105.321107
39. Shimbel A. Structural parameters of communication networks / A. Shimbel // Bulletin of Mathematical Biophysics. – 1953. – Vol. 15 (4). – P. 501–507. DOI: 10.1007/BF02476438
40. Dijkstra E. W. A note on two problems in connexion with graphs / E. W. Dijkstra // Numerische Mathematik. – 1959. – Vol. 1, Issue 1. – P. 269–271. DOI:10.1007/BF01386390.
41. Lipsky V. Combinatory for programmers / V. Lipsky. Moscow : Mir Press., 1988. – 213 p.
42. Fast Algorithm for Calculating Transitive Closures of Binary Relations in the Structure of a Network Object / [V. Batsamut, S. Manzura, O. Kosiak et al.] // International Journal of Computing. – 2021. – Vol. 20(4). – P. 560–566. DOI: 10.47839/ijc.20.4.2444

## ANALYSIS OF RISK TERMINAL FLOWS IN TECHNOGENIC SYSTEMS ARISING IN THE PROCESS OF THREAT IMPACT

**Sabat V. I.** – PhD, Associate Professor, Associate Professor of the Department of Information Multimedia Technologies of Ukrainian Academy of Printing, Lviv, Ukraine.

**Sikora L. S.** – Dr. Sc., Professor, Full Member of the Engineering Academy of Ukraine, Professor of the Department of Automated Control Systems of the Institute of Computer Sciences and Information Technologies, Lviv, Ukraine.

**Durnyak B. V.** – Dr. Sc., Professor, Honoured Worker of Science and Technology of Ukraine, Rector of Ukrainian Academy of Printing, Lviv, Ukraine.

**Povkhan I. F.** – Dr. Sc., Professor, Dean of the Faculty of Information Technologies, Uzhhorod National University, Uzhhorod, Ukraine.

**Polishchuk V. V.** – Dr. Sc., Associate Professor, Professor of the Department of Systems Software, Uzhhorod National University, Uzhgorod, Ukraine.

### ABSTRACT

**Context.** The analysis of the risk terminal flows in technogenic systems is carried out, which arise in the process of the impact of informational and cognitive threats in the automated document management system as part of the hierarchical production system.

**Objective.** The goal of this work is to develop a complex model for assessing the risk of the document management system failure as part of a hierarchical production system under the active threats.

The object of the research is the process of functioning of complex systems with a hierarchical structure, in which automated document management systems with a high level of data flow protection for decision-making are used to provide the information quality control of technological processes.

The subjects of the research are the methods and means of constructing an information protection system to ensure the reliable functioning of automated document management systems and making targeted decisions in hierarchical structures with minimal risk of exposure to external threats and attacks.

**Method.** For the first time, the cause-and-effect diagram of the event formation with the active action of threat factors and attacks is substantiated and developed, the interpretation of risk in a technogenic system is defined, and the risk in the space of states is presented as a change in the trajectory in the system transitions to the limit operation mode. For the first time, a category diagram of the structure of risk generation under the threat factors and a system-category diagram of interaction in the system  $risk \leftrightarrow emergency-active\ nature$  is constructed, a system-category scheme of risk formation under the active threat factors is suggested. For the first time, a cognitive diagram for assessing losses in the event of a risk situation arising from incorrect actions of the personnel is substantiated.

**Results.** As a result of the research, a system-category diagram of the impact of a set of threats on the system functioning mode and process is constructed, a method is developed for calculating the level of system strategic security of energy-active hierarchical systems in the process of attacks and threats, and a complex model for assessing the risk of a system functioning failure under active threats is suggested.

**Conclusions.** Under the action of active obstacles, cognitive and system factors at the operational and strategic levels of the control hierarchy, due to wrong decisions and informational disorientation, emergency situations and risks of the system function loss and its target-orientation arise. The analysis of a set of risks and the suggested category diagram of the risk generation structure under the impact of threat factors form the basis of the development of the probability structure of the risk concept based on the  $attack \leftrightarrow consequence$  model, as well as the construction of a system-category diagram of the interaction in the game  $active\ factor \leftrightarrow accident\ risk$ . This, in turn, makes it possible to construct a system-category scheme for the formation of risk terminal flows in technogenic systems that arise in the process of threat impact. A complex model for assessing the risk of system failure under threats can be used to construct protection systems for any hierarchical control structures of technogenic systems.

**KEYWORDS:** technogenic systems, threats, vulnerabilities, risk assessment, decision making, control of hierarchical systems.

### NOMENCLATURE

$\alpha_{risk}$  is an assessment of the risk level;  
 $\alpha_d$  is an acceptable risk level;  
 $KIA_i$  is a crisis information agent;  
 $F_i$  are attack factors;  
 $SV_i$  is a technogenic system (TS);  
 $\Pi_i$  are resource flows;  
 $\{x_i\}$  is a vector of possible alternatives on the set  $X$  of the target space partitioning;  
 $\omega_i$  are external negative factors from the set  $\Omega$ ;

$f$  is a function of the relationship between decision  $x$  and consequence  $y$ ;  
 $StratU$  is a control strategy;  
 $R_i$  are control system resources;  
 $V(x_i, \omega)$  is a selected alternative for the space partitioning of the system states;  
 $T_m$  is an interval of the threat;  
 $D_i$  is an event in the security system, situation;  
 $\tau_i$  is a time of the emergency event;  
 $C_i$  is a consequence of an emergency event;  
 $SitIIS$  is a situation in the space of states;

$Resurs(TO)$  is a resource of the technological object under the impact of external threat factors;

$R_d$  is a permissible resource value under normal conditions;

$Opt(U/S_i)$  is an optimal control strategy;

$V_i$  are losses upon successful completion of the attack;

$A(f_i)$  is an activity of the factor action on the aggregate structure and control system;

$P_i$  is a probability of the event occurrence;

$N_{di}^f$  is a consequence of the influencing factor  $f_i$ ;

$V_{ni}^f$  is a significance of the consequences after the effect of the influencing factor  $f_i$ ;

$U(W)$  is a utility function;

$Sh_W$  is a scale of the utility function with the values interval  $I$ ;

$K_i$  is a mechanisms and requirements for information confidentiality;

$\{Z_i\}$  are types of threats;

$\{Va_i\}$  are loss of information authenticity;

$r(\Pi\alpha_{risk})$  is a flow of cognitive risks;

$\Pi\mu(\alpha_{risk})$  are errors of the personnel;

$\alpha_{risk\ opt}$  is an optimal risk value, which does not exceed the permissible value.

## INTRODUCTION

The functioning of any technogenic system in the environment is vulnerable to a variety of interdependent negative factors, independent disturbances and threats that can lead to crisis situations, accidents and catastrophes, if preventive measures to assess the accident risk are not developed and decision-making procedures are not established for control under active informational and cognitive threats. Usually, such procedures and means of counteracting negative factors are provided at the beginning of the design of the protection system and are described in the form of provisions in the security policy of any organization with a hierarchical control system. However, there are negative factors and threats that may arise in the process of the system functioning, which also require the use of a probability approach to determine the risks associated with the action of active threats. Therefore, an important task of the research in the protection system is and will always be the analysis of risk flows, which together form a complex model for assessing the control failure risk for technogenic systems.

If one proceeds from the system concept of assessing the situation in local, district, regional infrastructures, it can be concluded that they were not ready for aggressive attacks on their structure and control process, because they were based on the concept of terminal stability. The action of informational and infrastructural attacks of an

aggressive type led to their collapse and the emergence of emergency situations with a high level of accident risk.

**The object of the research** is the process of functioning of complex systems with a hierarchical structure, in which automated document management systems with a high level of data flow protection for decision-making are used to provide the information quality control of technological processes.

**The subjects of the research** are the methods and means of constructing an information protection system to ensure the reliable functioning of automated document management systems and making targeted decisions in hierarchical structures with minimal risk of exposure to external threats and attacks.

**The goal of this work** is to develop a complex model for assessing the risk of the document management system failure as part of a hierarchical production system under active threats.

## 1 PROBLEM STATEMENT

To achieve the goal of the scientific research, it is necessary to solve the following tasks:

– for the first time, to develop a cause-and-effect diagram of the event formation during the active action of threat factors and attacks, to form the interpretation of risk in a technogenic system in the space of states as a change in the trajectory in the system transitions to the limit functioning mode;

– for the first time, to construct a category diagram of the structure of risk generation under the threat factors and a system-category diagram of interaction in the system  $risk \leftrightarrow emergency-active\ nature$ ;

– for the first time, to develop a system-category scheme of risk formation in the conditions of active threat factors and a cognitive diagram for assessing losses in the event of a risk situation arising from incorrect actions of personnel;

– to test and verify the suggested complex model for assessing the risk of the document management system failure as part of a hierarchical production system for the example of risk assessment of printing productions, as well as to offer a system-category interaction diagram in the game  $active\ factor \leftrightarrow accident\ risk$ .

Problem setting is formulated as follows. Let one have some research object  $SV_i$  in the input, which is assessed by many indicators depending on the control strategy  $StratU$  and security policy. The output may deviate from the control target due to the set of threats  $\{Z_i\}$  and the attack factors  $F_i$ , which these threats use. As a result of the construction of category diagrams of risk formation in the space of system states under the threat factors in the time interval  $T_m$ , it is possible to assess the amount of losses in the risk situations and in the event of incorrect actions of the personnel, the level of which can, for example, be represented by using the models of utility functions for risk assessment  $\alpha_{risk} = \{\alpha_{r1}(F_1), \dots, \alpha_{rn}(F_n)\}$  from the interval  $[0; 1]$ . Moreover, the indicators  $\alpha_{ri}(F_i)$

can represent the whole system of criteria and models, on the basis of which one aggregated assessment  $\alpha_{risk\ opt}$  on the scale of the utility function  $Sh_{\mathcal{W}}$  is derived for each factor of threats and attacks  $F_i$ , which is equated to the acceptable risk level  $\alpha_d$ .

In addition to quantitative assessments, the reasoning of experts analysing the object is used for the research object. For this, on the basis of experience and knowledge about the research object  $SV_i$ , a group of experts (or an expert) analyses it, draws conclusions and assigns one linguistic assessment to each indicator  $\alpha_{risk}$ , from the set  $r(\Pi\alpha_{risk}) = \{F_i; Z_i; KLA_i; Va_i; \mu(\Pi\alpha_{risk})\}$ .

Thus, for a complex assessment of the risk level for a technogenic system, it is necessary to conduct an analysis of the attack factors, multiple threats, loss of authenticity of the control information and possible errors of the personnel, which lead to risk situations and deviations of the system from target orientation. On the basis of the presented input data, for the research object  $SV_i$ , it is necessary to derive the initial aggregate assessment of the risk level in the process of the threats impact on the technogenic system  $\alpha_{risk} \in [0;1]$ . Analysing the value of the assessment  $\alpha_{risk}$  and equating it to the permissible values  $\alpha_d$  established at the design stage and prescribed in the security policy  $SV_i$ , it is possible to adjust the level of its protection by implementing optimal countermeasures.

## 2 REVIEW OF THE LITERATURE

Analysis of the problem of the emergency and risk situations occurrence, under active threats and attacks, shows the importance of constructing models of attack penetration channels and methods of action on system nodes. In the work of A. M. Shurygin [1], the use of statistical methods for forecasting risk situations and risks assessing is substantiated; the basic concepts, essence, objectives and methods of information protection and the organization of printing productions are revealed in the educational manuals of Vietnamese scientists [2, 3]; in the publications dedicated to the printing industry, the production of securities [4, 5], the peculiarities of the control of printing production and the protection of printing products are considered; scientific publications are devoted to various thorough means of information protection: by Schneier Bruce [6] on applied cryptography, by Mykhailo Sikorsky, and others [7] on analysis of malicious software; the work of L. H. Koval and other scientists [8] is on the analysis of biometric identification methods. With the help of the above-mentioned works, the methods and means of printing production protection are developed, also using the normative legal framework of the Laws of Ukraine [9], as well as the methods of software and hardware protection of network technologies [10], mobile communication devices [11].

In the work [12], modern concepts of information protection against information attacks and system threats are

substantiated; the authority control in information security systems, in particular, new approaches to the protection of printing companies, is discussed in the monograph [13].

In [14], information technologies for controlling complex hierarchical systems under threats and information attacks are studied. The concept of risk assessment is formed on the basis of determining the probability and frequency of threats and vulnerabilities for the company assets. The scientific article [15] describes an expert model for assessing risks and security incidents of airport network and information systems, based on intellectual analysis of knowledge using the apparatus of fuzzy sets. Nevertheless, the model is not able to assess the system impact on the process of the functioning system control. The document [16] analyses the concept of risk and safety of subway passengers in cases of malicious technogenic incidents. As a result, the importance of protecting passengers in terms of increasing safety and avoiding dangerous conditions is proven, using the example of the Athens metro system. These studies reveal the essence of hierarchical systems and their vulnerability to technogenic disasters under the external attacks and internal threats impact.

Modern developed methods of analysis of general industrial control systems of hierarchical technogenic structures are presented in the works of foreign scientists [17, 18], and they are given in [19, 20] specifically for complex systems.

The paper [21] presents the application of an object-oriented Bayesian network for scenario risk assessment. A model of probability coverage of key factors influencing accidents in fragmented structures is developed. In the studies [22], a model-based methodology for hybrid control of risk assessment of reliability, availability, maintainability, and safety for critically important systems is proposed. As a result, a method of cyber security risk analysis for industrial control systems is created. Agrawal et al. [23] define the ontology to represent the ISO/IEC 27,005, 2018 standards, with the aim of providing a step-by-step understanding of the meaning of security concepts and their relationships. For example, cyber security ontologies are developed by Arbanas & Čubrilo [24], who are able to construct 52 security ontologies. Researchers such as Blanco et al. [25] considered 31 security ontologies. Both studies group the security ontologies into three categories: general, specific, and theoretical.

Two popular risk assessment methods approved for the nuclear sphere use probability risk assessment [26, 27], and others use dynamic Bayesian networks [28, 29].

In the works of J. Rabcan and others [30, 31], the problem of developing a new algorithm with the application of a fuzzy classifier for signal classification is proposed. The results can be used to automate the process of constructing recognition models using precedents.

Nevertheless, a thorough systematic analysis of terminal risks in technogenic systems under the active threats of resource, information and cognitive types based on category models of influence channels on the control process has not been carried out to date.

In view of all the above-mentioned facts, it has been decided to carry out an innovative study on the development of a complex model for assessing the risk of the document management system failure as part of a hierarchical production system under the active threats impact.

### 3 MATERIALS AND METHODS

To analyse risks in technogenic systems and construct schemes and methods for their minimization and control, it is necessary to apply a risk analysis methodology based on the following components:

- the source of risk factors (activity, power, channels);
- the scenario of active actions and the factors impact on the system functioning process (structural, resource, information failure);
- the analysis of the results of the factors effects on the system;
- the structure of the intrusion zone of the active system, which allows the action of an intelligent threat agent through unidentified channels;
- active agents that form the means of the function and structure collapse of the technogenic hierarchical system at the physical level and errors in projects.

The source of risk is related to the consequences of active actions through the scenario – the chain of events of risk implementation in the system under certain conditions, which leads to negative consequences and accidents (Fig. 1).

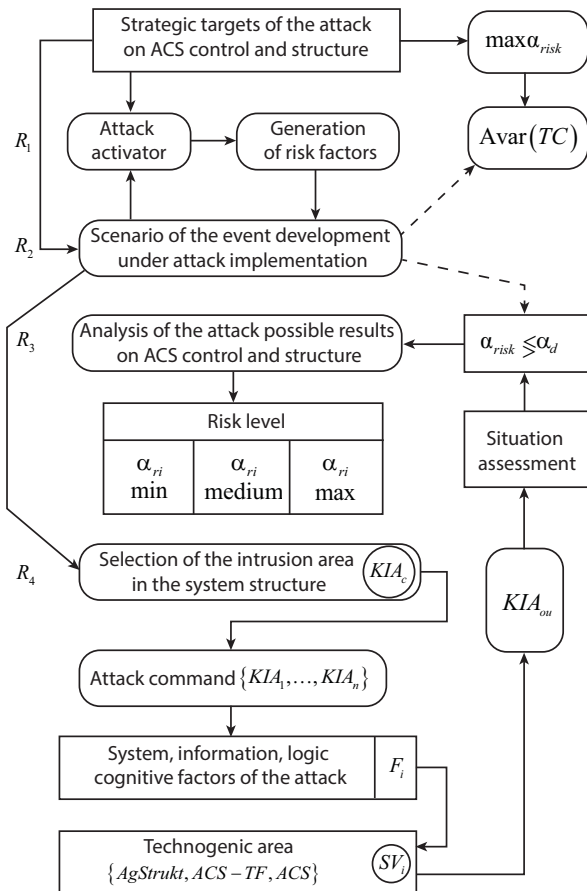


Figure 1 – Informational and cognitive map of the system intrusion process

Chains, paths and directions of connections are actually scenarios of the development of a dangerous situation from the point of view of different positions. They describe the scenarios of events that can happen to the system under the action of active factors generated by the source of risks – an active agent, an attacking system, a hidden internal crisis agent, errors of managerial personnel when making decisions.

With the action of active obstacles, cognitive and system factors at the operational and strategic levels of the control hierarchy, due to incorrect decisions and information disorientation, emergency situations and risks of loss of system functions and its target-orientation arise (Fig. 2).

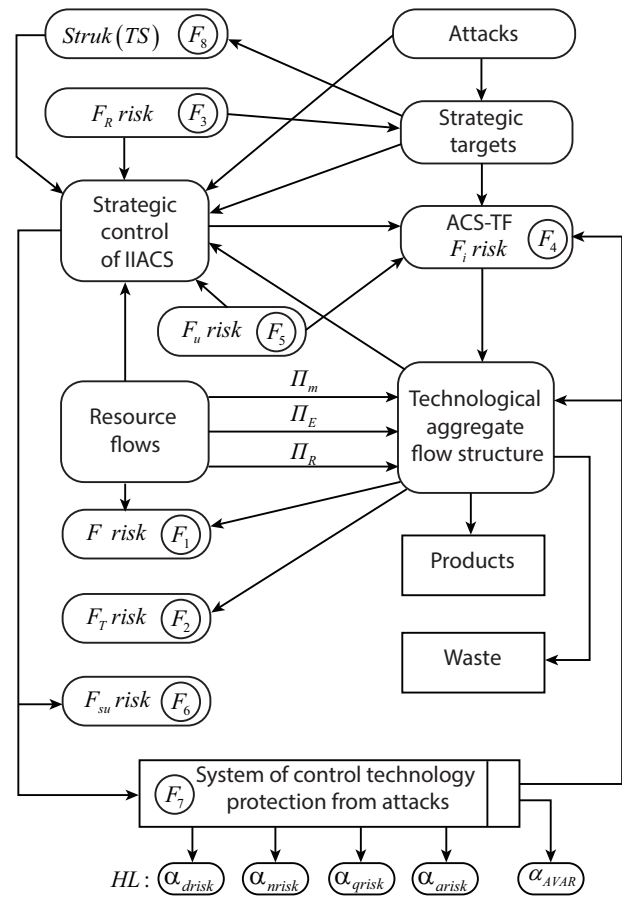


Figure 2 – Category diagram of the formation of hybrid complexes of active risks under the attack factors flow

Here are the schemes for forming a set of risks:

1. Physical risks – physical loss of resources during the system operation, collapse of structures, blocks, aggregates, nodes;
2. Technogenic risks – failure of systems, networks, computers, communications, DoS attacks and threatening destructive actions, power supply failures;
3. Position and cognitive risks – non-compliance with the criteria, regarding the position and abilities of the person, which leads to wrong decisions;

4. Information risks – loss of data, unauthorized access to ports, terminals, cryptosystems, attacks on databases and information security systems;

5. Management risks – access to decision-making systems, analysis, control, forgery of powers, disorientation of personnel;

6. General organizational risks – conflicts, personnel risks, erroneous setting of targets, inadequacy of situations perception in the control system, inadequacy of strategic targets, sabotage;

7. Risks of violation of the system security level – associated with attacks on the existing levels of the control system, selection and processing of data and decisions;

8. System risks – associated with possible errors when selecting a system concept (targets, structure, dynamics, data processing, control strategies) – by control and operational, project staff.

The level of risk (permissible, sufficient, limit, warning, emergency) is the basis of the classification of both system functioning modes and the assessment of the reliability of the functioning of aggregates, blocks, and control processes in complex integrated hierarchical systems.

The risk of control failure will be analysed in conditions of stochastic uncertainty of the situation under the action of active factors in the space of system states and the target of the control system.

Let one have  $\{x_i\}$  – a vector of possible alternatives on an admissible set  $X$  of the space partitioning of targets (modes, states).

A rational selection is made according to the consequences that the control action leads to under the factors impact  $\omega_i \in \Omega$  at the moment  $t_i$ .

The connection between the decision  $f$  regarding  $x$  and the consequence  $y$  is determined by:

$$\exists Strat(u/f) : y = f(x, \omega, t_i),$$

$$f : (X \times \Omega \times T) \rightarrow Y \rightarrow \langle f_{onm} \in F \rangle,$$

where  $\{f(x, \omega)/t_i\}$  – is the function that characterizes (costs-spending);  $f : (X \times \Omega) \rightarrow Y$  – is the model of the decision-making process;  $y = f(x_i |_{i=1}^n, \omega)$  – defines the process of calculating the consequences of the stochastic factor on  $x_i$ .

The two-stage decision-making process, under the influencing factor, has the following structure related to the cause-and-effect representation: [1]

1. The decision is made for the first move according to the alternative  $x_i \in X$ , then random factors are implemented  $\omega_i \in \Omega$  ( $AF_i \rightarrow \tau_i \rightarrow D_i \rightarrow Sit\Pi S$ ).

2. The decision is made  $y = Y(x_i, \omega_i)$ , which corrects  $x_i$  ( $\exists Strat U, U : x \rightarrow y$ ).

Implementation costs  $x_i$  will be  $f_1(x_i)$ , and for the implementation  $y = f_2(x, y, \omega)$ , thus, if:

$$\exists W(Resurs(TO) > R_d) \Rightarrow \exists Strat(U/f_1, f_2) :$$

$$\left\langle f(x_i, \omega) = \begin{cases} f_1(x_i), \\ f_2(x^*, y(x, \omega), \omega) \end{cases} \right\rangle \rightarrow \langle Opt(U/S_i) \rangle.$$

From the above, it is possible to construct loss functions according to the risk level of control selection on the terminal decision-making cycle for active target-oriented control. At the same time, the level of losses under decision-making risk may depend on the selection of strategies:

1.  $F_1(x) = f(x_i, \omega^*)$  – are normalized losses of resources, products;

2.  $F_2(x) = \max f(x, \omega)$  – are the highest losses of the system type;

3.  $F_3(x) = M_{\omega_i \in \Omega} f(x_i, \omega) = f(x, \omega)P(d\omega)$  – is the function of the probable average risk under the action of active threats flows;

4.  $F_4(x) = P\{f(x_i, \omega) \leq f(x, \omega) \int P(d\omega)\} : (\omega | f(x, \omega) \leq c)$  – is the

probability of losses under active threats that create an emergency situation.

The risk assessment is determined according to the situation and the activation of the threats, i.e.:

$\alpha_{risk}(\omega_i | x_i) = f(x_i, \omega) = F(x)$  – is determined by the selected alternative on the space partitioning of the system state in the form of a category diagram (Fig. 3).

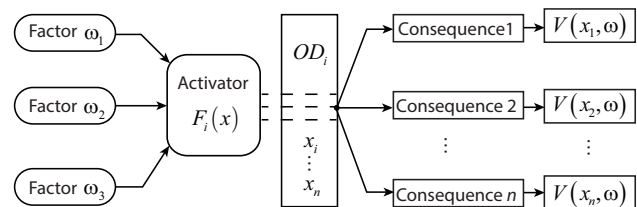


Figure 3 – Category diagram of the structure of risk generation under threat factors

According to the above category diagram, the probability structure of the concept of risk is justified based on the model  $\langle attack \rightarrow consequence \rangle$ :

$$\forall t_i \in T_m, \left\langle \begin{array}{l} \text{if } P_i(t_i) > 0, \text{ then} \\ Risk = \bigcup_{i=1}^n (P_i, C_i) \Rightarrow \max(C_r/V_i) \end{array} \right\rangle,$$

where  $P_i$  – is the event probability;  $C_i$  – is its consequence;  $V_i$  – are resource losses at  $C_r$ ;  $T_m$  – is the interval of the treat action.

Then the general risk form for a certain type of active influencing factor is determined according to the formula:

$$\forall t_i \in T_m, V_m^f > 0, P_i > 0: \left( \begin{array}{l} Risk = \bigcup_{i=1}^n (P_i, N_{di}^f, V_{ni}^f) \rightarrow \min \alpha_v, \\ \text{at } A(f_i) \rightarrow 0 \end{array} \right),$$

Then the structure of the interaction system  $\langle \text{structure} \leftrightarrow \text{active factors} \rangle$  can be presented in the form of a category diagram of threats impact in Fig. 4, 5, 6.

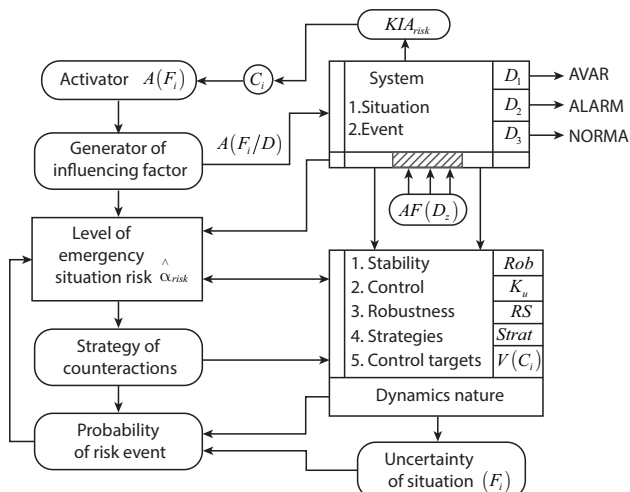


Figure 4 – System-category diagram of interaction in the game  $\langle \text{active factor} \leftrightarrow \text{accident risk} \rangle$

According to the above analysis of events in the system, which is affected by both control actions and active threats, a set of utility functions of the action of a complex of factors  $\{F_k |_{k=1,m}\}$  is constructed from the selected loss minimization strategies.

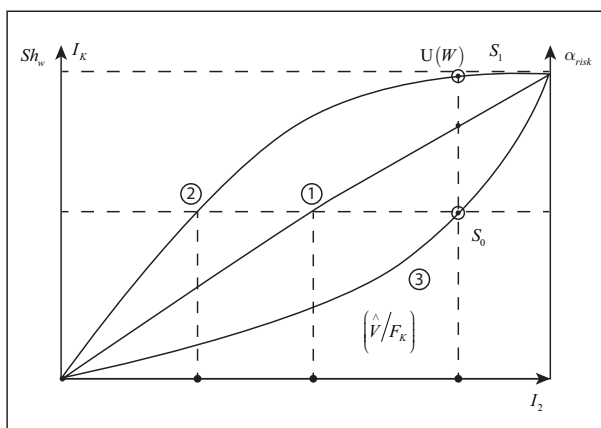


Figure 5 – Graph of the utility function when making decisions

Symbols in Fig. 5:  $Sh_w$  – is the scale of the utility function with the value interval  $I_k$ ;  $U(W)$  – is the Neumann-Morgenstern utility function when forming the decision selection on the set of alterna-

tives;  $\left( \hat{W} = \sum_{i=1}^n W_i P_i \right)$  – is the expected benefit of the event

with the probability of the consequences of the individual selection of the behaviour strategy (target-oriented) in relation to the system, situation  $(U(W) = P(U|S) - (1-p)U(S))$  – i.e. between maximum and minimum  $(S_1, S_0)$ .

Graphs of the utility function  $U(W)$  when making decisions of the PMD-KIA (the person making the decision), with the expected benefit  $\left( \hat{V}/F_{Ki} \right)$ , determines the

losses in case of incorrect actions of the operator with different types of behaviour:

1. PMD<sub>1</sub> – indifferent to risk (cognitively resistant);
2. PMD – not prone to risk (mentally unstable);
3. PMD<sub>3</sub> – prone to risk when making management decisions.

According to the determination of the usefulness level from control actions, under threats, a system-category scheme of the formation of a risk situation in the technogenic hierarchical structure (TS) is constructed (Fig. 6).

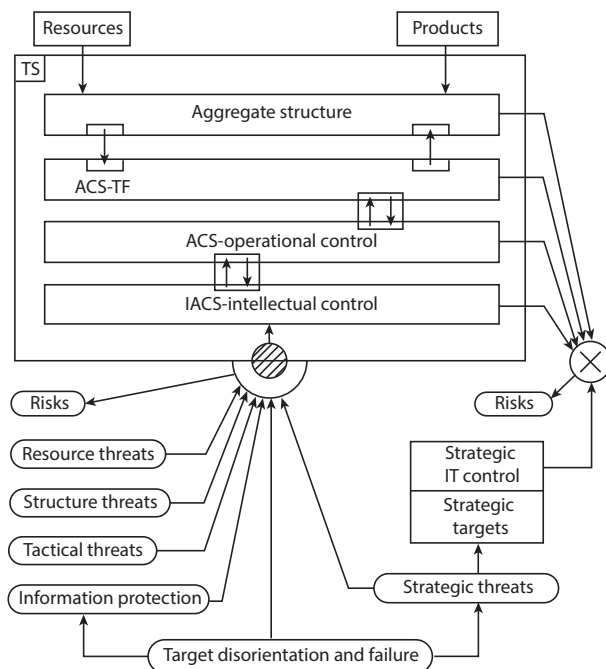


Figure 6 – System-category scheme of risk formation under active threats to the functioning of technogenic hierarchical structure objects

#### 4 EXPERIMENTS

Informational and cognitive threats will be analysed and studied in the automated document management system as part of the hierarchical production system.

The security system is primarily focused on identifying threats and, accordingly, risks of losses for the organization. Such threats include threats to confidentiality, integrity, availability, accountability, authenticity, and reliability of information.



When assessing the above threats, it is important to directly analyse the security issues with system assets, as this can affect possible threats and, as a result, the selection of protective measures. Usually, such an assessment contains a specific approach to each organization in particular, and also requires the involvement of specialists not only in the field of information protection, but also directly from the organization's specialists, who can assess the system assets, analyse the consequences of damage to these assets and means for restoring the working mode of the system after possible incidents. An information security incident is any unforeseen or undesirable event that can disrupt operations or information security. Information security incidents are: loss of services, equipment or devices; system failures or overloads; user errors; non-compliance with policies or guidelines; violation of physical protection measures; uncontrolled system changes; software failures and failures of technical means; violation of access rules.

Let one consider each of the threats in more detail, analyse the possible consequences of successfully implementing attacks on the system and the necessary countermeasures to reduce them [2–4], and construct a category diagram of the threats actions to the system and control processes (Fig. 7, 8).

Threats to the data privacy about the entire system usually use information access attacks. Threats to privacy can be countered by appropriate privacy and identification services. The methods of such attacks on the information system (*IS*) are usually reduced to three unauthorized actions: spying, eavesdropping and interception.

The mechanisms for ensuring the confidentiality of information in the form of files are presented in Table 1.

Table 1 – File confidentiality mechanisms and requirements for them

Mechanisms for ensuring confidentiality	$K_1$	Physical security control
	$K_2$	Access control to the computer files
	$K_3$	File encryption
File confidentiality requirements	$K_4$	Identification and authentication
	$K_5$	Correct setting of the computer system
	$K_6$	Correct key control when using encryption

Spying ( $ZA_1$ ) is carried out by accessing unauthorized information and viewing it. If these are paper documents, first of all, it is necessary to ensure their physical protection, preventing third parties from accessing confidential information (security, locks, safes, surveillance and alarm systems, etc.) [12]. If it is an electronic document, then in addition to physical protection and identification of authorized users, it is necessary to implement their authentication methods. At the same time, the following mechanisms of document files confidentiality and requirements for them should be taken into account [13]:

Eavesdropping ( $ZA_2$ ) can be carried out, for example, by connecting to a line and listening to a telephone con-

versation. Electromagnetic radiation from sources of information dissemination in *IS* can also be used.

Let one study the threats from information attacks.

If the information is transmitted through an internal or external network, it is possible to intercept it.

In the case of interception, it is important to take into account the fact that an attacker can carry out such an attack, but in order to prevent its further successful development, it is necessary to implement countermeasures to encrypt such information. At the same time, it is necessary to introduce reliable encryption technologies for information flows, or all communication traffic (Fig. 7) [6].

Loss of privacy can lead to the following negative consequences:

- loss of public trust or lowering of the organization's image in the society ( $VR_1$ );
- liability before the law, including liability for violation of legislation in the field of data protection ( $VR_2$ );
- negative influence on the organization policy ( $VR_3$ );
- creating a threat to the safety of the organization's personnel ( $VR_4$ );
- financial losses ( $VR_5$ ).

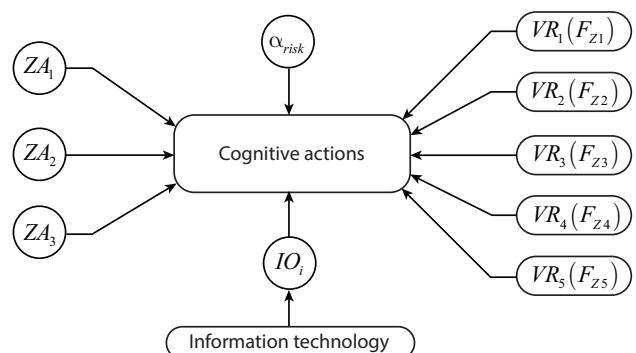


Figure 7 – Cognitive diagram for assessing losses in the risk situation

Symbols in Fig. 7:  $\{ZA_i\}$  – are types of threats affecting the system;  $\alpha_{risk}$  – is the risk level;  $IO_i$  – is a type of the information operation;  $\{VR_i\}$  – are losses under the factors impact  $\{F_i\}$ .

Let one consider informational threats to the data integrity as the basis for errors in the formation of management decisions.

The threat to the information integrity leads to the implementation of attacks on its modification, so all possible negative events that can lead to such incidents should also be analysed (Fig. 5.10). Such events include the following:

- physical access ( $ZB_1$ ) to the places of storage of information carriers – it threatens the integrity of the information located on such carriers);

- maintenance error ( $ZB_2$ ) – it occurs if it is carried out irregularly or without compliance with all security policy procedures, then the integrity of the relevant information is at risk;
- malicious code ( $ZB_3$ ) – it can lead to a violation of integrity, for example, if changes are made to data or files by an unauthorized person who gained access using malicious code, or such changes are made by the code itself [7];
- spoofing a legitimate user ID ( $ZB_4$ ) – it can be used to bypass authentication and related services and security functions. As a result, integrity problems may arise every time when the information is accessed and modified under the guise of a legitimate user [8];

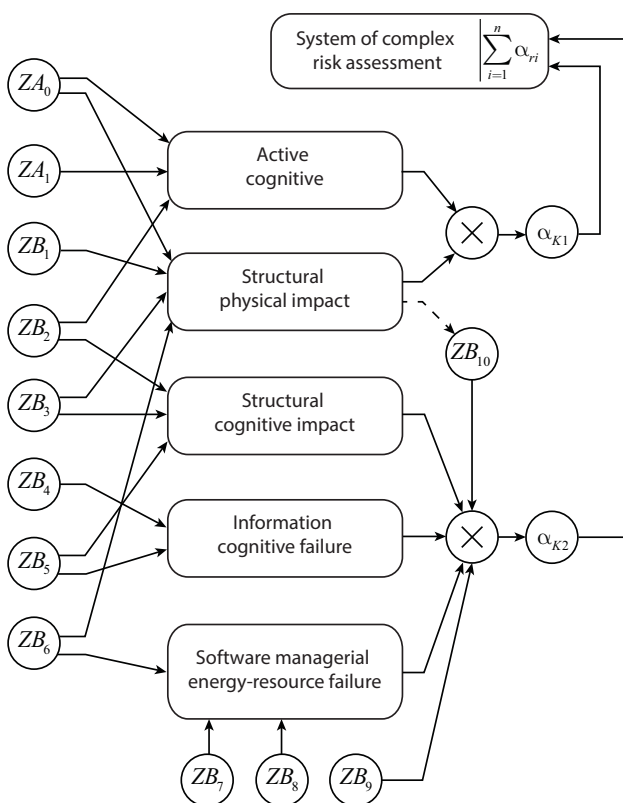


Figure 8 – System-category diagram of the impact of a set of threats on the system

- sending messages along a wrong or changed route ( $ZB_5$ ) – it can lead to a violation of integrity, for example, when messages are changed and then transmitted to the original addressee;
- software failure ( $ZB_6$ ) – it may violate the integrity of data and information that is processed with the help of such software;
- failures in power sources ( $ZB_7$ ) – (electrical supply and ventilation) – it can cause integrity problems, if such violations are the cause of other malfunctions. For example, power outages can lead to hardware failure, technical

malfunctions, or data storage problems. Such problems include technical malfunctions;

- unauthorized access to computers, data, services and applications ( $ZB_8$ ) – it can become a threat to the information integrity if their unauthorized change is possible [13];
- use of unauthorized programs and data ( $ZB_9$ ) – it creates a threat to the information integrity in the storage device and during its processing in the system, if these programs and data are used to illegally change information or contain malicious code;
- unauthorized access to the place of storage of information carriers ( $ZB_{10}$ ) – it may endanger the integrity of this information, because in this case, unauthorized changes to the information recorded on this information carrier are possible.

The integrity service monitors the reliability of information. With the proper level of organization, it gives users confidence that the information is correct and no one has changed it. The integrity service must work together with the identification service to perform a reliable verification of the authenticity of the person, his authenticity to the admission level. Therefore, the integrity service is a “shield” against modification attacks [18].

Loss of integrity can lead to the following:

- making wrong decisions ( $NR$ );
- failure in the organization’s commercial operations ( $NK_0$ );
- loss of public trust or lowering of the public image of the organization ( $VD_r$ ) that performs social activities;
- financial losses ( $VF$ ) from crisis situations and emergencies;
- liability before the law, including liability for violations of legislation in the field of data protection ( $ZV_Z$ ).

Hybrid threats to the availability of an automated hierarchical control system from attacking agents usually include events that allow attackers to carry out denial-of-service attacks. Among security specialists, such attacks are also called DoS (Denial-of-Service) attacks. The following threats are considered that can lead to the specified attacks [14]:

- destructive actions ( $ZRD_1$ ) – destructive attacks, which can also be called vandalism;
- physical access ( $ZRD_2$ ) – to storage locations of information carriers – it threatens the readiness for functioning of storage facilities;
- equipment malfunction ( $ZRD_3$ ) – connection and failure of communication services;
- maintenance error ( $ZDp$ ) – it often occurs if maintenance is carried out irregularly or with errors;
- malicious code ( $ZKd$ ) – it can be used to bypass the authentication and related services and security functions. As a result, this can lead to a loss of accessibility.

For example, if data or files are destroyed by a person who gained unauthorized access using malicious code, or the code itself erases files [7];

- spoofing a legitimate user ID ( $ZIK$ ) – it can be used to bypass the authentication and all related services and security functions. As a result, accessibility problems may arise every time when impersonating a legitimate user makes it possible to delete or destroy information [5];

- incorrect routing ( $ZMi$ ) or change of message routing [10];

- abuse of resources ( $ZR$ ), which leads to failures of the network mode.

- natural disasters ( $ZKS$ ) – impact on the structure and energy supply;

- software failures ( $ZZp$ ) – it can lead to the unavailability of data and information that is processed with the help of these programs;

- disruptions in supply ( $ZNRp$ ) – it can lead to availability problems if these disruptions are the cause of other malfunctions. For example, power outages can cause hardware failure, technical malfunctions, or data storage problems. Therefore, it is advisable to provide workplaces with uninterrupted power supply units [12];

- technical malfunctions ( $ZNI$ ) of nodes, blocks, system structures;

- theft ( $ZRs$ ) of spare sets for communication and control systems, which leads to an accident;

- traffic overload ( $ZNpt$ ) – it will reduce system reliability;

- transmission errors ( $ZNp$ ) – effects of interference on data transmission channels and systems;

- unauthorized access to computers ( $ZND$ ), data, services and applications – it can become a threat to the information accessibility, if unauthorized destruction of this information is possible;

- use of unauthorized programs and data ( $ZNpd$ ) – it creates a threat to the information accessibility in the storage device and during processing in the system, if programs and data are used to destroy information or if they contain malicious code;

- unauthorized access to storage locations of information carriers ( $ZNdn$ ) – it can lead to a risk of information accessibility, since in this case unauthorized destruction of information recorded on these carriers is possible [14].

In accordance with the presented conditions, a diagram of risk formation in the technogenic system is constructed (Fig. 9).

The information accessibility service supports its readiness for work, allows access to computer systems, data stored in these systems, and programs. This service provides the information transfer between two endpoints or computer systems. It is mainly about the information

presented in electronic form (but also suitable for ordinary documents).

Loss of access to data in ADMS and ACS can lead to the following consequences and the action of threat factors:

- making wrong decisions ( $F_{nr}$ );

- inability to perform important assigned tasks ( $F_{nc}$ );

- loss of public trust or lowering of the public image of the organization ( $F_{vd}$ );

- financial losses ( $F_{vf}$ );

- liability before the law, including liability for violations of legislation in the field of data protection and non-fulfilment of contracts within the established terms ( $F_{vz}$ );

- significant costs for restoring ( $F_{vv}$ ) the system structure, communication channels, and software.

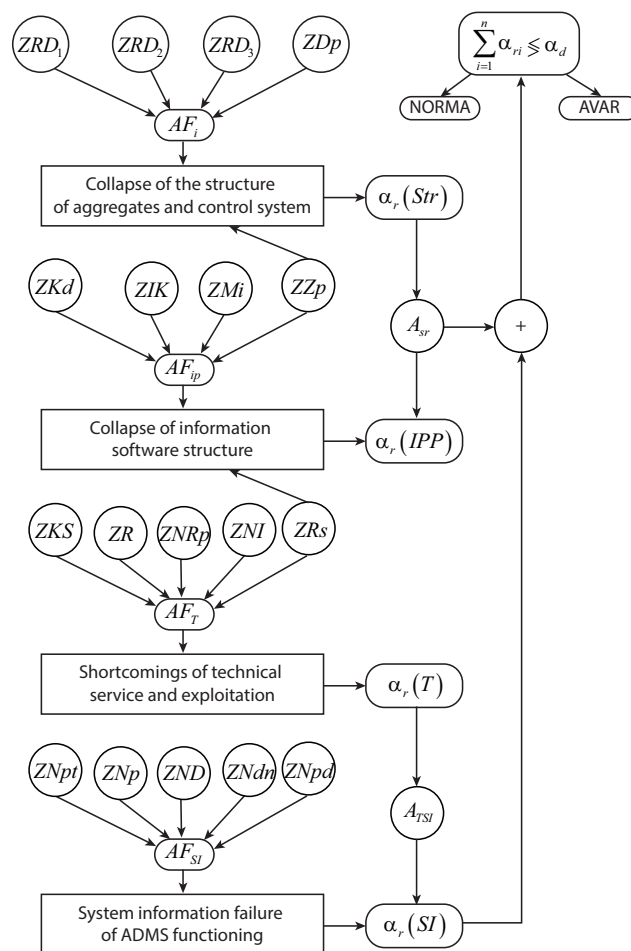


Figure 9 – Diagram of formation of the system accident risk under a complex of threats

Let one consider the threats of accountability for the hierarchy levels of the cyber technogenic system in the implementation of targeted control tasks.

When protecting accountability, any threat that may lead to the performance of actions that are not characteristic of this object or entity should be taken into account: collective use of accounts; lack of possibility of operational control of actions; imitation of a legitimate user (masquerade); software failure; unauthorized access to the computer, data, services and applications; unsatisfactory authentication. Such threats typically use disclaimer attacks [12, 14].

The loss of accountability, under the factors impact of the system functioning process failure can lead to the following informational and cognitive consequences:

- manipulation of the system by users ( $ZF_m$ );
- deception of personnel at the levels of the system hierarchy (disinformation) ( $ZF$ );
- industrial spying (possibility of attacks) ( $ZF_p$ );
- uncontrolled actions leading to emergency situations ( $ZF_d$ );
- false accusations of incorrect decisions of individuals ( $ZF_z$ );
- liability before the law, including liability for violation of legislation in the field of data protection ( $ZF_v$ ).

Let one consider threats to authenticity in the event of data failure and system disorientation.

Trust in authenticity can be undermined by any threat that causes a person, system, or process to doubt that an object is who it claims to be. Examples of the occurrence of such a situation are the change of data without proper control, the origin of unverified or unsupported data (Fig. 10). [14]

Loss of authenticity can lead to the following consequences in the ADMS system:

- deception of the lower levels ( $Va_{1,1}$ ) and disorientation of the upper levels of the control hierarchy ( $Va_{1,2}$ );
- use of reliable processes with unreliable data (which can lead to a misleading result) ( $Va_2$ );
- manipulation of the organization from the outside – structure ( $Va_{3,1}$ ) and targets ( $Va_{3,2}$ );
- industrial spying regarding target documents ( $Va_4$ );
- false accusations that lead to conflicts in the system ( $Va_5$ );
- liability before the law, including liability for violation of national legislation in the field of data protection ( $Va_6$ ).

According to these factors, category diagrams of the formation of hybrid attacks on the information authenticity are developed for each control system (Fig. 10).

Threats of management data inaccuracies are not necessary for assessing the situation and making control decisions.

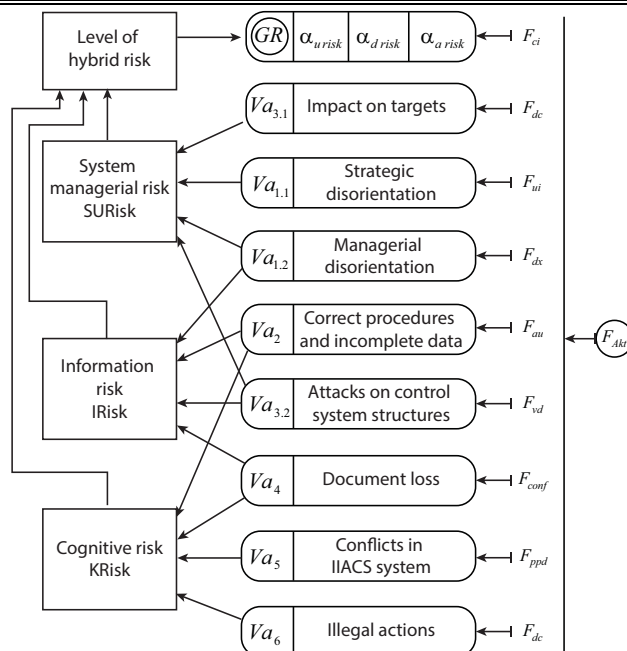


Figure 10 – Category diagrams of the authenticity loss of the control process under internal and external cognitive threats

Any threat that can lead to inconsistent behaviour of systems or processes leads to a decrease in reliability. Examples of such threats are illogical functioning of the system and unreliable suppliers. Decreasing credibility leads to poor customer service and loss of trust.

The loss of reliability can lead to the following consequences in the process of control decision-making under the threat factors:

- deception of the personnel, which leads to a conflict;
- loss of market share due to disorganization;
- decrease in motivation in the work of the organization’s personnel, which leads to the emergence of risk situations during control;
- unreliability of suppliers;
- decrease in customer confidence in the service system;
- liability before the law, including liability for violations of legislation in the field of data protection.

Any attack is implemented through the performance of certain actions that disrupt the performance of the protection system and the automated control system as a whole. For a successful attack, an attacker needs to identify a weak spot in the chain of the protection system and, due to the threats and vulnerabilities of the system, make an illegal intervention in its work. At the same time, the main attention of attackers is aimed at security services, which are focused on countering attacks. Therefore, for the reliable functioning of the security system, an important task is to establish the operation of all its security services and analyse the formation of the characteristics of influencing factors and risk components (Fig. 11).

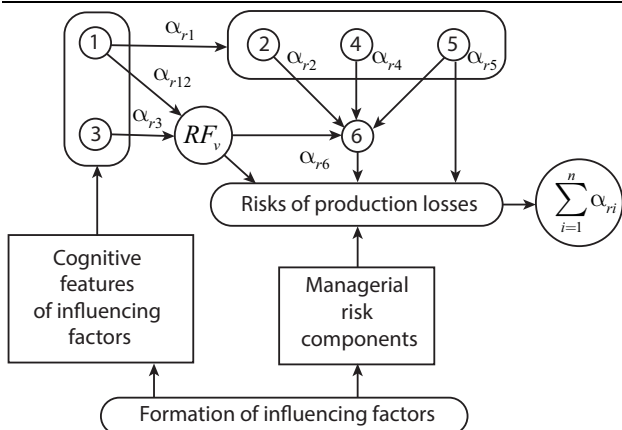


Figure 11 – A complex model for assessing the system failure risk under active threats

### 5 RESULTS

The analysis of literary sources and the results of research on the behaviour and knowledge level of operational personnel in energy (KMDandA, ACS) at thermal power plants (coal) was carried out on the basis of a system analysis of cognitive engineering psychology, methods of the theory of knowledge. This became the basis for the development of a table of crisis skills, which provide the possibility of selection and formation of operational anti-crisis management teams in the conditions of active threats and attacks of a complex type (Table 2).

Table 2 – Factors and skills

№	Requirements to the activity	Factor	Coefficient $\alpha_r$
1.	Information processing of images	$FI_v$	0.1–0.5
2.	Operational actions	$FI_{od}$	0.1–0.95
3.	To form images of situations	$FI_{syt}$	0.1–0.5
4.	Factor of target-oriented actions	$FCS_u$	0.05–0.95
5.	Factor of action tactics generation	$FG_{id}$	0.1–0.35
6.	Factor of sensory information perception	$FSS_i$	0.1–0.25
7.	Factor of skills to implement strategy	$FR_{str}$	0.1–0.9
8.	Ability to master knowledge	$KFIZ_1$	0.1–0.5
9.	Ability to construct models of objects and event scenarios	$KFIZ_2$	0.05–0.3
10.	System target-orientation when the mode is broken	$KIZ_3$	0.1–0.95
11.	Formation of images of terminal situations	$KIZ_4$	0.05–0.25
12.	Analysis of the dynamics of events in the system, control modes	$KIZ_5$	0.05–0.5
13.	Forecast of the consequences of a person's managerial actions	$KIZ_6$	0.05–0.95
14.	Genetic features of a person's thinking	$IKK_g$	0.01–0.3
15.	Motivational and will-power ability to make decisions	$IKK_M$	0.5–0.95
16.	Cognitive stress resistance	$SKI$	0.1–0.95
17.	The level of system and professional knowledge of the operator	$RSP_I$	0.5–0.9
18.	Ability to apply knowledge in crisis situations (creativity)	$FKSit$	0.1–0.95

Tables of this type ensure the construction of effective tests for the selection of personnel for teams of operational and strategic control levels, which are capable of resisting active threats of a high-risk level.

According to Table 2, a category diagram of influencing factors on the formation of possible cognitive risks – staff errors is constructed (Fig. 12).

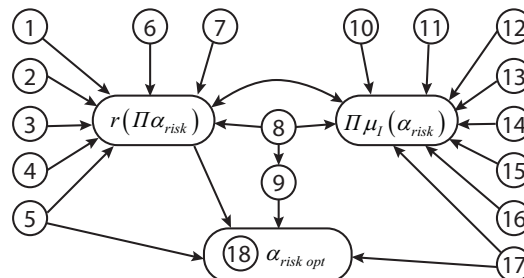


Figure 12 – Category diagram of cognitive risk assessment

### 6 DISCUSSION

The developed complex model of risk assessment increases the accuracy and objectivity of assessing risk situations, because on the one hand it uses quantitative assessments of the research object (based on data on possible losses in the event of incidents) according to various models and criteria, and on the other hand it uses the experience, knowledge and competencies of experts in the subject area.

The general concept of this approach can be applied not only to companies of the printing industry, but also to any technogenic structures with a hierarchical control structure. Quantitative assessment of losses is presented in the model input obtained according to the analysis of possible threats and vulnerabilities, for example, organizational assets, cognitive characteristics of factors affecting the technogenic structure, taking into account risk coefficients, characteristics of various management components of risk and linguistic considerations of experts-specialists in the field of security and control and decision-making systems. The risks of production losses are formed in the terminal cycle of production and can also be changed under the active threats in the process of production and control. It is necessary to take into account all the threats and vulnerabilities of such technogenic hierarchical systems, and only then it is possible to determine a complex indicator of the system failure risk under active threats. The weight of each constituent indicator of risk can be determined on the basis of the probability approach of the occurrence of one or another event depending on the level of threat to a certain asset, or on the basis of statistical surveys of experts in a specific subject area of research. As a rule, probability approaches coincide with practical statistical results, but unforeseen force majeure circumstances may arise, which, for example, we are currently experiencing during military operations, which destroy all scientific forecasts and conclusions of experts. Assessment and risk analysis of technogenic hierarchical structures was carried out in peacetime, during 2020–2021, for which there are relevant Implementa-

tion Acts, in particular for the printing production. Based on the received complex assessment of the system malfunction risk, recommendations can be provided for improving the protection system for a certain production and its separate structures, which increases the reliability of the security system.

The developed model of a complex assessment of the system failure risk under active threats does not give a quantitative assessment of the risk, as it determines weak points for the system functioning. In particular, the amount of cognitive risk is difficult to quantify until an incident occurs – a staff or manager error. Therefore, the study substantiated a cognitive diagram for assessing losses in the event of a risk situation with incorrect actions of personnel. The graph of the utility function in decision-making of a decision-maker with expected gains and losses, in case of incorrect actions with different types of behaviour, which affects the cognitive component of the risk level, is presented. In the process of testing with the help of a category diagram *active factor* ↔ *accident risk*, it is possible to reduce the level of cognitive risk, thanks to the increase in the competencies and skills of the personnel.

Obtaining a quantitative risk assessment for the research object has a number of advantages, namely: it combines quantitative (reliable) assessments with the experience, knowledge, and competencies of experts in the subject area; it is based on the definition of the decision-making utility function; the managerial personnel can be trained and decision-making levels can be adjusted according to the method proposed in the research. The suggested category diagrams and a complex risk assessment model can be used not only in the process of designing protection systems, but also in the process of system operation to solve those problems where there is no data for training and where it is necessary to periodically monitor the security system.

The disadvantages of this approach include the fact that the obtained coefficients of cognitive risk depend on the division of the interval  $[0,1]$ , and their values depend on the competencies of experts in certain subject areas of research. The training of managerial personnel in the security system also depends on this.

## CONCLUSIONS

The paper solves the scientific and applied task of developing a complex model for assessing the risk of system failure under active threats, which, on the one hand, uses quantitative assessments of the object, and on the other hand, takes into account the experience, knowledge and competencies of experts in the relevant subject area.

**The scientific novelty** of the conducted research is as follows:

– for the first time, a cause-and-effect diagram of the event formation during the active action of threat factors and attacks has been developed;

– the interpretation of risk in a technogenic system in the space of states as a change in the trajectory in the

system transitions to the limit functioning mode has been improved;

– for the first time, a category diagram of the structure of risk generation under the impact of threat factors and a system-category diagram of interaction in the system *risk* ↔ *emergency-active nature* have been constructed;

– for the first time, a system-category scheme of risk formation in the conditions of active threat factors and a cognitive diagram for assessing losses in the event of a risk situation arising from incorrect actions of personnel have been developed;

– the proposed complex model for assessing the risk of the document management system failure as part of a hierarchical production system for the example of risk assessment of printing productions has been tested and verified, and in addition, a system-category interaction diagram in the game *active factor* ↔ *accident risk* has been suggested.

**The practical significance** of the obtained results is that the proposed model of complex risk assessment has been tested in the document management system as part of the hierarchical system of printing production and can be used in various technogenic hierarchical systems when solving managerial decision-making tasks, designing and improving protection systems.

**The further research** of the problem can be seen in the development of software for assessing the risk of system functioning under active threats to technogenic hierarchical structures.

“shield”

## ACKNOWLEDGEMENTS

The work has been performed within the framework of the state-budget research themes of Ukrainian Academy of Printing: “Development of information technologies for the protection of electronic resources in automated document management systems”; Uzhhorod National University: “Methods and means of software engineering for implementing big data analytics processes based on information and technical platforms of electronic science”.

The authors would also like to express their gratitude their partners Aquaflex Plus LLC, Prime-Pak LLC, Sofi Company LLC for the data provided for the implementation of the experimental part of the study.

## REFERENCES

1. Shurygin A. M. Applied stochastics: robustness, estimation, forecast. Moscow, Finance and statistics, 2000, 224 p.
2. Kavun S. V., Nosov V. V., Manzhai O. V. Information security. Tutorial. Kharkiv, PH. KhNEU, 2008, 352 p.
3. Veretilnyk T. I., Mysnyk L. D., Mysnyk B. V., Kapitani R. B. Organization of publishing and printing activities: Tutorial Cherkasy. Cherkasy, State Technology University, 2020, 157 p. [Electronic resource] <https://er.chdtu.edu.ua/bitstream/ChSTU/3380/1/ORGANIZATION%20POLIGRAPHIC%20ACTIVITY.pdf>
4. Kovaleva V. V., Samarin Yu. N. Selection of management system for a printing company, *CompuArt. Journal for printers and publishers*, 2007, No. 11, pp. 61–64.

5. Honcharov S. V. Financial security of the securities market of Ukraine. Poltava, Poltava State Agrarian Academy, 2019, pp. 40–42.
6. Schneier Bruce. Applied cryptography. Protocols, algorithms, source texts in C language. 2nd edition. Moscow, Triumph, 2002, 816 p.
7. Michael S., Andrew H. Practical Malware Analysis: The Hands – On Guide to Dissecting Malicious Software; translated from English. Chernikov S., St. Petersburg, 2018, 786 p.
8. Koval L. H., Zlepko S. M., Novitskyi H. M., Krekoten E. H. Methods and technologies of biometric identification according to the results of literary sources, *Scientific notes of TNU named after V.I. Vernadskyi*. Vinnytsia, VNTU, 2019, Vol. 30 (69), Part 1, No. 2, pp. 104–112. [Electronic resource] [https://www.tech.vernadskyjournals.in.ua/journals/2019/2\\_2\\_019/part\\_1/19.pdf](https://www.tech.vernadskyjournals.in.ua/journals/2019/2_2_019/part_1/19.pdf).
9. Law of Ukraine “On electronic digital signature”, *Bulletin of the Verkhovna Rada*, 2003, No. 36, P. 276.
10. Schneider B. Secrets and Lies: Digital Security in a Networked World. New-York, WCP, 2002, 368 p.
11. Senkivskyi V. M., Petyak Y. F., Kozak R. O., Lytovchenko O. V. Information technology for effective data protection of publishing systems on mobile devices. Lviv, UAP, 2020, 272 p.
12. Bobalo Y. Ya., Horbaty I. V., Bondarev A. P. Information security. Lviv, Lviv Polytechnic University, 2019, 580 p.
13. Durnyak B. V., Sabat V. I., Shvedova L. E. Authority control in information protection systems. Lviv, UAP, 2016, 148 p.
14. Sabat V. Sikora L., Durnyak B., Lysa N., Fedevych O. Information technologies of active control of complex hierarchical systems under threats and information attacks, *The 3rd International Workshop on Intelligent Information Technologies & Systems of Information Security (IntellTISIS-2022)*. Khmelnytskyi, Ukraine, May 25–27, 2022. <https://ceur-ws.org/Vol-3156/paper23.pdf>
15. Kelemen M., Polishchuk V., Gavurová B., Andoga R., Szabo S., Yang W., Christodoulakis J., Gera M., Kozuba J., Kaľavský P., Antoško M. Educational Model for Evaluation of Airport NIS Security for Safe and Sustainable Air Transport. *Sustainability*, 2020, 12, 6352. <https://doi.org/10.3390/su12166352>.
16. Milioti Christina, Kepaptsoglou Konstantinos, Deloukas Alexandros, Apostolopoulou Efthymia Valuation of man-made incident risk perception in public transport: The case of the Athens metro, *International Journal of Transportation Science and Technology*, 2022, Vol. 11, pp. 578–588. <https://doi.org/10.1016/j.ijst.2021.07.003>.
17. Sicard F., Zamai É., Flaus J. M. An approach based on behavioral models and critical states distance notion for improving cybersecurity of industrial control systems, *Reliab Eng Syst Saf*, 2019, Vol. 188, pp. 584–603. 10.1016/J.RESS.2019.03.020
18. Cormier A., Ng C. Integrating cybersecurity in hazard and risk analyses, *J Loss Prev Process Ind*, 2020, Vol. 64. Article 104044, 10.1016/j.jlp.2020.104044
19. Schmittner C., Gruber T., Puschner P., Schoitsch E. Security application of Failure Mode and Effect Analysis (FMEA), *Computer safety, reliability, and security*. Springer International Publishing, Cham, 2014, pp. 310–325.
20. Vessels L., Heffner K., Johnson D. Cybersecurity risk assessment for space systems, *2019 IEEE Space Comput Conf. (SCC)*, 2019, pp. 11–19. 10.1109/SpaceComp.2019.00006
21. Domeh Vindex, Obeng Francis, Khan Faisal, Bose Neil, Sanli Elizabeth Risk analysis of man overboard scenario in a small fishing vessel, *Ocean Engineering*, 2021, Vol. 229, Article 108979. <https://doi.org/10.1016/j.oceaneng.2021.108979>.
22. Alanen Jarmo, Linnosmaa Joonas, Malm Timo, Papakonstantinou Nikolaos, Ahonen Toni, Heikkilä Eetu, Tiusanen Risto Hybrid ontology for safety, security, and dependability risk assessments and Security Threat Analysis (STA) method for industrial control systems, *Reliability Engineering & System Safety*, 2022, Vol. 220, Article 108270. <https://doi.org/10.1016/j.res.2021.108270>.
23. Agrawal V. A. Comparative study on information security risk analysis methods, *J Comput (Taipei)*, 2017, pp. 57–67. 10.17706/jcp.12.1.57-67
24. Arbanas K., Čubrilo M. Ontology in information security, *J Inf Org Sci*, 2015, Vol. 39, pp. 107–136.
25. Blanco C. Lasheras J., Fernández-Medina E., Valencia-García R., Toval A. Basis for an integrated security ontology according to a systematic review of existing proposals, *Comput Stand Interfaces*, 2011, Vol. 33, pp. 372–388.
26. Zhou T., Modarres M., Droguet E. L. Multi-unit nuclear power plant probabilistic risk assessment: a comprehensive survey, *Reliab Eng Syst Saf*, 2021, Vol. 213. Article 107782. 10.1016/J.RESS.2021.107782
27. Modarres M., Zhou T., Massoud M. Advances in multi-unit nuclear power plant probabilistic risk assessment, *Reliab Eng Syst Saf*, 2017, Vol. 157, pp. 87–100. 10.1016/J.RESS.2016.08.005
28. Kim J., Shah A.U.A., Kang H.G. Dynamic risk assessment with bayesian network and clustering analysis, *Reliab Eng Syst Saf*, 2020, Vol. 201, Article 106959, 10.1016/J.RESS.2020.106959
29. DeJesus Segarra J., Bensi M., Modarres M. A bayesian network approach for modeling dependent seismic failures in a nuclear power plant probabilistic risk assessment, *Reliab Eng Syst Saf*, 2021, Vol. 213, Article 107678. 10.1016/J.RESS.2021.107678
30. Rabcan J., Levashenko V., Zaitseva E., Kvassay M., Subbotin S. Application of Fuzzy Decision Tree for Signal Classification, *IEEE Transactions on Industrial*, 2019, No. 15(10), pp. 5425–5434. <https://doi.org/10.1109/TII.2019.2904845>
31. Rabcan J., Levashenko V., Zaitseva E., Kvassay M., Subbotin S. Non-destructive diagnostic of aircraft engine blades by Fuzzy Decision Tree, *Engineering Structures*, 2019, No. 197, P. 109396. <https://doi.org/10.1016/j.engstruct.2019.109396>

Received 16.01.2023.  
Accepted 01.02.2023.

## АНАЛІЗ ТЕРМІНАЛЬНИХ ПОТОКІВ РИЗИКІВ У ТЕХНОГЕННИХ СИСТЕМАХ, ЯКІ ВИНИКАЮТЬ В ПРОЦЕСІ ВПЛИВУ ЗАГРОЗ

**Сабат В. І.** – канд. техн. наук, доцент, доцент кафедри інформаційних мультимедійних технологій Української академії друкарства, Львів, Україна.

**Сікора Л. С.** – д-р техн. наук, професор, дійсний член Інженерної Академії України, професор кафедри автоматизованих систем управління Інституту комп'ютерних наук та інформаційних технологій, Львів, Україна.

**Дурняк Б. В.** – д-р. техн. наук, професор, заслужений діяч науки і техніки України, ректор Української академії друкарства, Львів, Україна.

**Повхан І.Ф.** – д-р техн. наук, професор, декан факультету інформаційних технологій ДВНЗ «Ужгородський національний університет», м. Ужгород, Україна.

**Поліщук В.В.** – д-р техн. наук, доцент, професор кафедри програмного забезпечення систем ДВНЗ «Ужгородський національний університет», Ужгород, Україна.

### АНОТАЦІЯ

**Актуальність.** Проведено аналіз термінальних потоків ризиків в техногенних системах, які виникають в процесі впливу інформаційних і когнітивних загроз в автоматизованій системі управління та документообігу в складі ієрархічної системи виробництва.

**Метою** даної роботи є розроблення комплексної моделі оцінки ризику збою системи управління та документообігу в складі ієрархічної системи виробництва при дії активних загроз.

Об'єктом дослідження є процес функціонування складних систем з ієрархічною структурою, в яких для інформаційного забезпечення якісного управління технологічними процесами використовуються автоматизовані системи документообігу з високим рівнем захисту потоків даних для прийняття рішень.

Предметом дослідження є методи та засоби побудови системи захисту інформації для забезпечення надійного функціонування автоматизованих систем документообігу та прийняття цільових рішень в ієрархічних структурах з мінімальним ризиком впливу зовнішніх загроз і атак.

**Метод.** Вперше обґрунтовано і розроблено причинно-наслідкову діаграму формування події при активній дії факторів загроз і атак, визначено трактування ризику в техногенній системі та представлено ризик у просторі станів як зміну траскторії при переході системи в граничний режим функціонування. Вперше побудовано категорну діаграму структури породження ризиків при дії факторів загроз та системно-категорну діаграму взаємодії в системі *ризик ↔ аварійно-активний характер*, запропоновано системно-категорну схему формування ризиків в умовах дії активних факторів загроз. Вперше обґрунтовано когнітивну діаграму для оцінки втрат при виникненні ризикової ситуації при некоректних діях персоналу.

**Результати.** В результаті досліджень побудовано системно-категорну діаграму впливу комплексу загроз на режим і процес функціонування системи, розроблено метод обчислення рівня системної стратегічної безпеки енергоактивних ієрархічних систем в процесі дії атак і загроз та запропоновано комплексну модель оцінки ризику збою функціонування системи при дії активних загроз.

**Висновки.** При дії активних завад, когнітивних і системних факторів на оперативному та стратегічному рівнях ієрархії управління із-за неправильних рішень та інформаційної дезорієнтації виникають аварійні ситуації та ризики втрати функцій системи і її цілеорієнтованості. Аналіз комплексу ризиків і запропонована категорна діаграма структури породження ризиків при дії факторів загроз, лягли в основу розроблення ймовірнісної структури поняття ризику на підставі моделі *атака ↔ наслідок*, а також побудови системно-категорної діаграми взаємодії в трі *активний фактор ↔ ризик аварії*. Це, в свою чергу, дало можливість побудови системно-категорної схеми формування термінальних потоків ризиків в техногенних системах, які виникають в процесі впливу загроз. Комплексна модель оцінки ризику збою системи при дії загроз може бути використана для побудови систем захисту для будь-яких ієрархічних структур управління техногенними системами.

**КЛЮЧОВІ СЛОВА:** техногенні системи, загрози, вразливості, оцінка ризику, прийняття рішень, управління ієрархічними системами.

### ЛІТЕРАТУРА

1. Shurygin A. M. Applied stochastics: robustness, estimation, forecast / A. M. Shurygin. – М. : Finance and statistics, 2000. – 224 p.
2. Kavun S. V. Information security. Tutorial / S. V. Kavun, V. V. Nosov, O. V. Manzhai. – Kharkiv : PH. KhNEU, 2008. – 352 p.
3. Organization of publishing and printing activities: Tutorial Cherkasy / [T. I. Veretilnyk, L. D. Mysnyk, B. V. Mysnyk, R. B. Kapitan]. – Cherkasy : State Technology University, 2020. – 157 p. [Electronic resource] <https://er.chdtu.edu.ua/bitstream/ChSTU/3380/1/ORGANIZATION%20POLIGRAPHIC%20ACTIVITY.pdf>
4. Kovaleva V. V. Selection of management system for a printing company / V. V. Kovaleva, Yu. N. Samarin // Compu-Art. Journal for printers and publishers. – 2007. – No. 11. – P. 61–64.
5. Honcharov S. V. Financial security of the securities market of Ukraine / S. V. Honcharov. – Poltava : Poltava State Agrarian Academy, 2019. – P. 40–42.
6. Schneier Bruce. Applied cryptography. Protocols, algorithms, source texts in C language. 2nd edition. – М. : Triumf, 2002. – 816 p.
7. Michael S. Practical Malware Analysis: The Hands – On Guide to Dissecting Malicious Software / S. Michael, H. Andrew ; translated from English. Chernikov S. – St. Petersburg, 2018. – 786 p.
8. Koval L. H. Methods and technologies of biometric identification according to the results of literary sources / [L. H. Koval, S. M. Zlepko, H. M. Novitskyi, E. H. Krekoten] // Scientific notes of TNU named after V. I. Vernadskyi. – Vinnytsia : VNTU, 2019. – Vol. 30 (69) – Part 1. – No. 2. – P. 104–112. [Electronic resource]



- [https://www.tech.vernadskyjournals.in.ua/journals/2019/2\\_2\\_019/part\\_1/19.pdf](https://www.tech.vernadskyjournals.in.ua/journals/2019/2_2_019/part_1/19.pdf).
9. Law of Ukraine “On electronic digital signature” // Bulletin of the Verkhovna Rada, 2003. – No. 36. – P. 276.
  10. Schneider B. Secrets and Lies: Digital Security in a Networked World / B. Schneider. – New-York : WCP, 2002. – 368 p.
  11. Information technology for effective data protection of publishing systems on mobile devices / [V. M. Senkivskiy, Y. F. Petyak, R. O. Kozak, O. V. Lytovchenko]. – Lviv : UAP, 2020. – 272 p.
  12. Bobalo Y. Ya. Information security / Y. Ya. Bobalo, I. V. Horbaty, A. P. Bondarev/ – Lviv : Lviv Polytechnic University, 2019. – 580 p.
  13. Durnyak B. V. Authority control in information protection systems / B. V. Durnyak, V. I. Sabat, L. E. Shvedova. – Lviv : UAP, 2016. – 148 p.
  14. Information technologies of active control of complex hierarchical systems under threats and information attacks / [V. Sabat, L. Sikora, B. Durnyak et al.] // The 3rd International Workshop on Intelligent Information Technologies & Systems of Information Security (IntellTISIS-2022) Khmelnytskyi, Ukraine, May 25–27, 2022. <https://ceur-ws.org/Vol-3156/paper23.pdf>
  15. Kelemen, M. Educational Model for Evaluation of Airport NIS Security for Safe and Sustainable Air Transport. Sustainability / [M. Kelemen, V. Polishchuk, B. Gavurová et al.]. – 2020. – Vol. 12. – P. 6352. <https://doi.org/10.3390/su12166352>.
  16. Valuation of man-made incident risk perception in public transport: The case of the Athens metro / [Christina Milioti, Konstantinos Kepaptsoglou, Alexandros Deloukas, Efthymia Apostolopoulou] // International Journal of Transportation Science and Technology. – 2022. – Vol. 11. – P. 578–588. <https://doi.org/10.1016/j.ijst.2021.07.003>.
  17. Sicard F. An approach based on behavioral models and critical states distance notion for improving cybersecurity of industrial control systems / F. Sicard, É. Zamai, J. M. Flaus // Reliab Eng Syst Saf, 2019. – Vol. 188. – P. 584–603. [10.1016/J.RESS.2019.03.020](https://doi.org/10.1016/J.RESS.2019.03.020)
  18. Cormier A. Integrating cybersecurity in hazard and risk analyses / A. Cormier, C. Ng. // J Loss Prev Process Ind, 2020. – Vol. 64. – Article 104044, [10.1016/j.jlp.2020.104044](https://doi.org/10.1016/j.jlp.2020.104044)
  19. Security application of Failure Mode and Effect Analysis (FMEA) / [C. Schmittner, T. Gruber, P. Puschner, E. Schoitsch] // Computer safety, reliability, and security. – Springer International Publishing, Cham, 2014. – P. 310–325.
  20. Vessels L. Cybersecurity risk assessment for space systems / K. Heffner, D. Johnson // 2019 IEEE Space Comput Conf. (SCC), 2019. – P. 11–19. [10.1109/SpaceComp.2019.00006](https://doi.org/10.1109/SpaceComp.2019.00006)
  21. Domeh Vindex. Risk analysis of man overboard scenario in a small fishing vessel / [Vindex Domeh, Francis Obeng, Faisal Khan et al.] // Ocean Engineering. – 2021. – Vol. 229. – Article 108979. <https://doi.org/10.1016/j.oceaneng.2021.108979>.
  22. Hybrid ontology for safety, security, and dependability risk assessments and Security Threat Analysis (STA) method for industrial control systems / [Jarmo Alanen, Joonas Linnosmaa, Timo Malm et al.] // Reliability Engineering & System Safety. – 2022. – Vol. 220. – Article 108270. <https://doi.org/10.1016/j.ress.2021.108270>.
  23. Agrawal V. A. Comparative study on information security risk analysis methods / V. A. Agrawal // J Comput (Taipei). – 2017. – P. 57–67. [10.17706/jcp.12.1.57-67](https://doi.org/10.17706/jcp.12.1.57-67)
  24. Arbanas K. Ontology in information security / K. Arbanas, M. Čubrilo // J Inf Org Sci, 2015. – Vol. 39. – P. 107–136.
  25. Basis for an integrated security ontology according to a systematic review of existing proposals / [C. Blanco, J. Lasheras, E. Fernández-Medina et al.] // Comput Stand Interfaces. – 2011. – Vol. 33 – P. 372–388.
  26. Zhou T. Multi-unit nuclear power plant probabilistic risk assessment: a comprehensive survey / T. Zhou, M. Modarres, E. L. Droguett // Reliab Eng Syst Saf. – 2021. – Vol. 213. – Article 107782. [10.1016/J.RESS.2021.107782](https://doi.org/10.1016/J.RESS.2021.107782)
  27. Modarres M. Advances in multi-unit nuclear power plant probabilistic risk assessment / M. Modarres, T. Zhou, M. Massoud // Reliab Eng Syst Saf. – 2017. – Vol. 157. – P. 87–100. [10.1016/J.RESS.2016.08.005](https://doi.org/10.1016/J.RESS.2016.08.005)
  28. Kim J. Dynamic risk assessment with bayesian network and clustering analysis / J. Kim, A.U.A. Shah, H.G. Kang // Reliab Eng Syst Saf. – 2020. – Vol. 201. – Article 106959, [10.1016/J.RESS.2020.106959](https://doi.org/10.1016/J.RESS.2020.106959)
  29. DeJesus Segarra J. A bayesian network approach for modeling dependent seismic failures in a nuclear power plant probabilistic risk assessment / J. DeJesus Segarra, M. Bensi, M. Modarres // Reliab Eng Syst Saf. – 2021. – Vol. 213 – Article 107678. [10.1016/J.RESS.2021.107678](https://doi.org/10.1016/J.RESS.2021.107678)
  30. Application of Fuzzy Decision Tree for Signal Classification // [J. Rabcan, V. Levashenko, E. Zaitseva et al.] // IEEE Transactions on Industrial. – 2019. – No. 15(10). – P. 5425–5434. <https://doi.org/10.1109/TII.2019.2904845>
  31. Non-destructive diagnostic of aircraft engine blades by Fuzzy Decision Tree // [Rabcan J., Levashenko V., Zaitseva E. et al.] // Engineering Structures. – 2019. – No. 197. – P. 109396. <https://doi.org/10.1016/j.engstruct.2019.109396>

*Наукове видання*

**Радіоелектроніка,  
інформатика,  
управління**

№ 1/2023

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 25.11.2022 Формат 60×84/8.  
Папір офс. Різогр. друк. Ум. друк. арк. 19,76.  
Тираж 300 прим. Зам. № 877

69063, м. Запоріжжя, НУ «Запорізька політехніка», друкарня, вул. Жуковського, 64

Свідоцтво суб'єкта видавничої справи  
ДК № 6952 від 22.10.2019.