Odesa Polytechnic National University

INFORMATION PROCESSING IN CONTROL AND DECISION-MAKING SYSTEMS. PROBLEMS AND SOLUTIONS.

MONOGRAPH

ODESA 2023

УДК 004:37:001:62
ББК 74.5(0)я431+74.6(0)я431+32.81(0)я431
И74

Рекомендовано до друку вченою порадою Національного університету «Одеська політехніка» (протокол №2 від 27.09.2023 р.)

Рецензенти:
Prof. Ye Zhengmao, Southern University (United States);
Prof. Algirdas Pakštas, British Computer Society (United Kingdom);
Prof. Anatoly Kargin, Ukrainian State University of Railway Transport (Ukraine);
Prof. Malakhov Evgeniy, Odessa National University, I.I. Mechnikov (Ukraine)

Авторський колектив:
Axak N., Bobok I., Dong C., Khlamov S., Kirchu P., Kobozeva A., Kondratenko G., Kondratenko Y., Kostiuk S., Krasnobayev V., Kushnaryov M., Kutiepov V., Kovalchuk D., Kovalchuk D.,Kozlov O.,Lysenko O.,Martynenko A.,Ponomarenko S.,Potienko O., Prots'ko I., Rudnichenko M., Rykmas R., Shibaeva N., Shevchenko A., Shpinareva I., Sitnikov T., Skakodub O., Slyusar V., Sushyn I., Tachinina O., Tatarnykov A., Otradska T., Petrov I., Vakulenko M., Vychuzhanin O., Vychuzhanin V., Yanko A., Zheng Y., Zhukov Y.

И74 Обробка **інформації** в системах управління та прийняття рішень. Проблеми та рішення. Монографія / Аксак Н., Бобок І. и др.; під наук. ред. проф. В.Вичужаніна – Одеса: НУ «ОМА», 2023 – 358 с.
Укр., англ. мовами.
ISBN 978-617-7857-33-3

У монографії відображені результати наукових досліджень у галузі інформаційних, інтелектуальних систем та аналізу даних, моделювання та розробки програм. Матеріали монографії будуть корисними для аспірантів, магістрантів, викладачів вищих навчальних закладів, що спеціалізуються на галузі ІТ-технологій.

УДК 004:37:001:62

ISBN 978-617-7857-33-3

INFORMATION PROCESSING IN CONTROL AND DECISION-MAKING SYSTEMS. PROBLEMS AND SOLUTIONS

CONTENTS

SECTION 10. INFORMATION SYSTEMS

PREFACE

The collective monograph presents the results of scientific research in the field of information systems and technologies, intelligent systems, data analysis, modeling and program development.

The monograph is compiled in the form of scientific articles-sections corresponding to thematic areas, which thoroughly reflect the results of research in the following areas: intelligent control technologies; control systems in robotic systems; information control systems; data security and cryptography; data mining technologies and big data; intellectual models and knowledge engineering technologies; mathematical and simulation modeling; intelligent systems and data analysis; multi-agent systems and distributed computing; control information system;information systems.

In the collective monograph, the authors pay attention to solving problems in the field of information systems and technologies using a fuzzy automatic control system for tracking mobile robotic platforms, an approach to determining the influence of the parameters of the spatial movement of a UAV on its characteristics. steganalytic method of declaring lsb attachments in digital video, sequences of digital images, the main directions of implementing the artificial intelligence strategy in Ukraine, intelligent monitoring of the technical condition of complex systems.

Considerable attention is paid to expanding neural network models with adaptive activation functions, the steganalytic method of declaring lsb-embeddings in digital video, the sequence of digital images, and intelligent monitoring of the technical condition of complex systems.

Through the research findings presented, readers will gain a useful body of knowledge needed to better understand problems and solutions in the field of information systems and technology.

The articles presented in the monograph correspond to the original author. Only the authors are responsible for the content of articles.

The materials of the monograph will be useful for graduate students, undergraduates, teachers of higher educational institutions specializing in the field of information systems and technologies.

Scientists consisting of 14 doctors of science, 14 candidates of science, and 12 applicants took part in the work on the collective monograph.

Section 1. Intelligent control of systems and technologies

UDC 004.8

MAIN DIRECTIONS FOR IMPLEMENTATION OF THE ARTIFICIAL INTELLIGENCE STRATEGY IN UKRAINE

Corr. Member of NAS of Ukraine, Dr.Sci. Shevchenko A.[1], Dr.Sci. Kondratenko Y.[1,2], Dr.Sci. Slyusar V.[1], Dr.Sci. Zhukov Y.[3], Ph.D. Kondratenko G.[2], Ph.D. Vakulenko M.[1]

[1]Institute of Artificial Intelligence Problems under the Ministry of Education and Science and National Academy of Sciences of Ukraine, Ukraine,
[2]Petro Mohyla Black Sea National University, Ukraine,
[3]C-Job Nikolayev, 17/6, Artyleriyska Str., Mykolaiv, 54006, Ukraine
ORCID: 0000-0002-0095-538X, 0000-0001-7736-883X, 0000-0002-2912-3149, 0000-0002-6391-4382, 0000-0002-8446-509, 0000-0003-0772-7950
[1] E-mail: rector_iai@ukr.net, ypk@ipai.net.ua, swadim@ukr.net, maxvakul@gmail.com
[2] E-mail: y_kondrat2002@yahoo.com, halyna.kondratenko@chmnu.edu.ua
[3]E-mail: y.zhukov@c-job.com.ua

Abstract. *This paper is devoted to the analysis of the specific focuses, directions, and peculiarities of the Strategy of Artificial Intelligence (AI) Development in Ukraine. The main paper's components are an analysis of the current state of the justification, development, and governmental approval of the National Strategy of AI in Ukraine; key elements and main priority areas of AI implementation according to IAIP-project "Strategy for AI Development in Ukraine"; proposals for AI development in short- and long-term perspectives and features of the AI implementation in Ukraine during the current wartime. Special attention is paid to such focuses in AI research and development as (a) the design of AI systems based on cognitive and conscience conceptions; (b) new solutions in intelligent robotic systems for ground, underwater and aerial applications; (c) AI perspectives in the marine industry; (d) prospective AI implementation in education; (e) linguistic competency of AI systems. The obtained, by the authors, results can be used for the development of strategic steps and plans in AI research and implementation on the governmental level of decision-making processes.*

Keywords: *Strategy, artificial intelligence, development, implementation, Ukraine, priorities, peculiarities, analysis, IAIP-project*

1. Introduction

Artificial intelligence (AI) plays a more and more important role in the different fields of human activity. Scientists and experts are expecting revolutionary results with AI development and implementation in medicine and healthcare, transportation, science, education, military and defense, manufacturing, agriculture, space exploration, and different services [1,2 3, 4, 5]. The new developments in the AI field are changing quickly and AI implementation areas are extending quickly. A new type of society is in the process of its establishment (Society 5.0), its chains of production, logistics, and social infrastructure will be based on artificial intelligence. The governments of developed countries understand the necessity of funding AI research for providing significant economic growth and for the leading position in the world's GDP competition. Many countries created their own Strategy for AI development and determined the priority areas for AI implementation, taking into account the features of their own economic situation, national interests, the indicators and possibilities in science, the level of the education system, and others. Among the countries with their own AI strategies are Canada, Japan, China, the United States, Brazil, Australia, Austria, Germany, and others. According to IQ-Holon publication [6] the governments of fifty countries from different continents have created and approved AI strategies in different forms and styles as plans, conceptions, roadmaps, extended and detailed strategies, executive orders, etc. Modern AI products have increasing implementation for solving different complex tasks with access for users based on fee or non-fee financial approaches. In particular, ChatGPT, GPT-4 and other AI platforms are very popular and very important with their huge potentials and possibilities [7, 8] for generating and correcting texts, consulting people in various spheres of human activities, reviewing and analyzing articles and reports, translating and calculating, transforming mathematical tasks, etc. At the same time, the powerful development and implementation of AI products led to many changes in the traditional styles of human life concerning changes in the labor market, in the set of personal and professional skills, in education processes (school and university curricula), and other changes. Many scientists, experts, policymakers, and entrepreneurs also widely discuss and focus on ethical issues in the AI design processes, the balance between the advantages and disadvantages of AI applications [9], and the dangers of AI implementation in powerful weapons, where AI will independently decide the fate of people. This paper aims to the analysis of the main focuses and features of the strategy for AI development in Ukraine. It is very important for consolidation and concentration of the research efforts for implementing AI in priority areas. The rest of this paper is organized as follows. Section 2 presents the developed “Strategy for AI development in Ukraine” with an analysis of its key components, Ukraine’s priorities in AI development, and specific features of AI implementation in the current wartime. In section 3, the authors discuss the approach to the design of AI systems based on conscience conception. New solutions in intelligent robotics for ground, underwater and aerial applications are considered in section 4. Sections 5 discuss the prospective AI implementations in the marine industry and section 6 – in education. Section 7 is devoted to the linguistic competency of AI systems. The paper ends with a conclusion in Section 8.

2. Strategy for AI development in Ukraine

Let us analyze the current state of the justification, development, and governmental approval of the “Strategy for Artificial Intelligence Development in Ukraine” based on the IAIP’s project on AI Strategy [10, 11] that is created under the leadership of the Institute of Artificial Intelligence Problems of the Ministry of Education and Science and National Academy of Sciences of Ukraine.

2.1. National AI development strategy in Ukraine: current state

The AI field is developing and implementing very fast in Ukraine. There are more than 2,000 software development companies in Ukraine specializing in the AI industry. Ukraine has made a progressive step in the publishing open data direction, especially during the past few years. Concerning the Global Open Data Index, Ukraine places 31st position in the world. The National Academy of Sciences, the Ministry of Digital Transformation, the Ministry of Education and Sciences, the Ministry of Strategic Industries, and many other governmental organizations in Ukraine are involved in the process of creating a National Strategy for AI Development and Implementation in Ukraine. As a result, the Conception for AI Development and Implementation was created in Ukraine and on 2 December 2020 [12] was approved by the Cabinet of Ministers of Ukraine.

In 2020 also was started the process of creating a detailed Strategy for the development of AI in Ukraine. The Institute of Artificial Intelligence Problems (IAIP) under the National Academy of Sciences (NASU) and the

Ministry of Education and Science of Ukraine (MESU) became the leading organization in the IAIP-project "Creating Strategy for the Development of AI in Ukraine" [10, 11]. Many Ukrainian scientists, who have scientific and practical experience in the AI field (including authors), were united in one team for creating, discussing, and promoting the Ukrainian AI Strategy. The main steps for the development of the Ukrainian AI strategy were defined by the next tasks: a) analysis and comparative review of the published national strategies of AI development in different countries from different continents; b) formation of a generalized presentation of the analytical AI centers' activities; c) determination of promising directions for developing AI in Ukraine; d) generalization of the basic terminology definitions, organizational principles, and main focuses of further research of Ukrainian scientists in the AI field; e) identification of the priority domains for implementation of advanced AI in Ukraine; f) formation of a list of necessary legislative, organizational, and investment measures for the implementation of the identified directions for the development of AI in Ukraine. The IAIP-project was successfully executed but, unfortunately, the Russian aggression on Ukraine in February 2022 seriously influenced the global discussions and final approval of this AI Strategy as National AI Strategy at the governmental level. Let us focus on the key components of the developed AI Strategy and the main priorities in the implementation of AI in Ukraine according to IAIP-project "Strategy for Artificial Intelligence Development in Ukraine".

2.2. IAIP-project of strategy for AI development in Ukraine: key content components and main priorities in AI implementation

The key content components in ten sections of the developed "Strategy for AI development in Ukraine" consist of an introduction and paradigm; basic AI concepts, definitions, and research directions; aims and objectives of the Ukrainian strategy for AI development; regulatory framework and current state of AI development and implementation in Ukraine; priority areas in Ukrainian economy for AI applications; scientific support, staffing, and funding for the national AI ecosystem; and evaluating the effectiveness of the Strategy for AI development in Ukraine. This AI strategy was created based on the Ukrainian national characteristics and interests, the necessity to extend AI research, and the implementation of the recent AI tool developments in different fields of the Ukrainian economy. During the process of the AI Strategy creation, IAIP sent letters of inquiry to over 300 different organizations, in particular, to the majority of ministries of Ukraine, scientific institutions, state and private institutions of higher education, and commercial organizations to determine the need to implement and use AI in their work. As a result, the next main priority areas for the implementation of AI in Ukraine were included in the Strategy for AI development in Ukraine" with detailed justifications and descriptions: AI in the National Security and Military-Industrial Complex of Ukraine; AI in Science and Education; AI in Medicine and Healthcare; AI in the Manufacturing Industry and Power Sector; AI in the Telecom Industry; AI in Transportation and Infrastructure; AI in Agriculture; AI in Ecology.

Besides, the priority areas for AI implementation in Ukraine according to [12] are economics, cybersecurity, information security, public administration, justice, legal regulation and ethics. The Strategy of AI Development in Ukraine (AIDU Strategy) [11] is designed for the period of 2023-2030, and its adoption process consists of two stages: (a) for 2023-2025; (b) for 2026-2030. To successfully implement the AIDU Strategy, the following immediate steps should be executed: Step 1. Approve and adopt the regulatory framework. Step 2. Create the supervisory board to monitor and accomplish the tasks declared in the AIDU Strategy. Step 3. Determine the roadmap of the AIDU Strategy; Step 4. Prioritize the objectives of the AIDU Strategy; Step 5. Accomplish the most prioritized and fundamental tasks; Step 6. Provide mechanisms for quarterly and annual control over the implementation of the AIDU Strategy (reporting, optional examination, etc.). Step 7. The final step is the reassessment of the AIDU Strategy, its analysis of compliance with the actualities of 2025, and, if necessary, its effective modification.

The AIDU Strategy should be supplemented with additional midterm (annual) deadlines, before which the aim and objectives of the relevant block must be completely accomplished. Each midterm period should be completed with an analytical report followed by an adjustment of the dynamic schedule. This component acts as a stimulus that will positively affect the intensity of the AIDU Strategy implementation. To effectively implement the AIDU Strategy, it is necessary to take the following measures by 2025:

– Create a regulatory framework that provides for the protection of economic and scientific data, as well as its storage in Ukraine.

– Provide scientific and theoretical support for the execution of the AIDU Strategy.

– Attract financial resources for the development of AI in Ukraine.

– Provide support for fundamental and applied scientific AI research.

– Increase the number of qualified AI employees and raise new technology awareness.

– Improve the digital literacy of the Ukrainian people.

– Build a national database system.

The main mechanism for the Strategy of Artificial Intelligence Development in Ukraine implementation is the annual action plans developed by the Committee on the Development and Implementation of Artificial Intelligence and approved by the Cabinet of Ministers of Ukraine.

2.3. Peculiarities of the AI implementation during the current wartime

The war in Ukraine has become the first high-tech war in human history, in which both sides of the conflict began utilizing the capabilities of so-called computational artificial intelligence (AI). The implementation of AI in Ukraine during wartime is characterized by its widespread use across various domains. First and foremost, AI plays a crucial role in tactical combat actions and military operations, particularly in enhancing the effectiveness of mass deployment of unmanned aerial vehicles (UAVs) for surveillance and reconnaissance tasks, and the evaluation of artillery fire effectiveness. According to experts, the deployment of UAVs accounts for over 70% of targets destroyed during combat operations. Developers quickly transitioned from using classical convolutional neural networks to segmenting objects based on various U-Net and PSPNet structures [13, 14]. In relatively simple object classification tasks, transfer learning and zero-shot learning methods based on neural networks previously trained on the ImageNet dataset [15, 16] performed well. Later, in a short period, the process of adapting known neural network technology, Object Detection (OD) [17], to field datasets gained popularity. This was used not only for automatic target detection but also for simultaneous classification under varying seasonal conditions. The most ambitious projects aim to implement object identification and target class recognition, for example, combat vehicles, tanks, logistical transport, etc. This can be extended to classify a specific type of object similarly. To solve all these mass deployment tasks, various versions of YOLO family neural networks were widely used [18]. Primarily, their operation takes place not onboard the UAVs but in the command post equipment. Importantly, object detection in images is combined with video tracking algorithms for real-time incoming video streams from onboard or stationary cameras of different spectral ranges in different domains. For example, Fig. 1 shows a fragment of video tracking based on YOLO5 Small of a high-speed motor boat. The neural network effectively tracks the boat and allows counting the number of people on board. Similar results of automatic detection and tracking of a drone by a neural network are shown in Fig. 2. This OD technology is also used for detecting unexploded ordnance on the seabed using underwater drones and assessing housing and infrastructure damage. An additional direction to enhance the capabilities of military information support was the application of intelligent chatbots in Telegram channels or based on separate mobile applications. These allow for alerts about the appearance of enemy machinery, means of air attacks, and so on. The chatbot boom has also covered areas such as psychological support for service members and legal assistance. Natural language processing (NLP) is generally a promising direction in the field of AI, especially considering the capabilities of the language model GPT-4 and its less powerful counterparts. With Ukraine receiving Western weapons samples, effective combat operation requires translations of NATO standards and corresponding technical documentation from various European languages into Ukrainian. In this regard, smartphone translators with built-in audio and optical text recognition feature from Google, as well as translation functions implemented in ChatGPT, have become handy. Furthermore, relying on local GPT-4 analogs such as LLaMA [19], Alpaca [20], etc., automatic analysis of combat reports from units can be provided. This allows for the prompt provision of information about the current battlefield situation to commanders upon their requests, facilitating rapid response to critical threats and decision-making support.

Fig.1. YOLO5 Small for the detection and tracking of a moving motorboat and people

Apart from the application of large language models (LLM), NLP algorithms, and neural networks for video tracking and image processing, an important direction is neural network processing of time series. This allows for predicting meteorological data for high-precision artillery firing, expenditure and needs in various resources, the evolution of satellite navigation correction adjustments over time, etc. Implementing AI in war conditions has its challenges. One of these is the necessity to ensure data security and protection against cyber-attacks. Given the heightened risk of cyber threats, artificial intelligence should be viewed as a potent player in cybersecurity efforts. It is employed in algorithms designed to detect and neutralize threats, as well as to protect critical infrastructure. The role of AI also extends to the coordination of humanitarian aid. In the logistics sector, AI ensures the efficient distribution and optimal delivery of assistance to those who need it most. In the realm of information warfare, artificial intelligence plays a significant role in detecting and countering disinformation campaigns. It enables the analysis of large volumes of data to discern patterns and trends in disinformation. Machine learning methods have become the de facto standard approach when performing social media and mass media publication analysis for Open Source Intelligence interests. The predictive analytics capabilities of AI are also employed to forecast enemy movements and the tactics of deploying weapons and military equipment.

a) b)

Fig. 2. Neural network tracking of UAVs in the evening (a, b)

This assists military planners in strategizing their actions and responding to potential threats. In addition, AI is utilized in the management and servicing of critical infrastructure during warfare. It aids in monitoring and predicting potential infrastructure failures, and coordinates repair and maintenance efforts. This information is based on the latest data available as of June 2023 and is continually supplemented with new evidence of the growing role of AI in all spheres of society amidst military operations.

3. Design of AI systems based on conscience conceptions

The novel authors' proposal in AI Strategy for Ukraine deals with the development, design and implementation of disruptive AI systems based on conscience conceptions. As its human cognate, artificial consciousness (AC) is a necessary attribute of an artificial personality with AI. Artificial consciousness manifests itself as an emergent global self-organized information phenomenon that evaluates and controls core processes of the system, exchanges data between system components to coordinate their behavior, provides for the social and personal perception of the environment, and conditions internal integration and external separation of the system [11]. It has been proposed that the AC modeling should include two sides of the same process: (i) the modeling of an attention schema as a mechanism of information selection and broadcasting; (ii)the modeling of the mechanism of information flow correlation. Successfully designing the AC is a complicated and multidisciplinary task but its solving will provide AI, which is friendly for humans. This approach suggests the synergetic treatment of AC that may be represented by a strange attractor and correspondingly simulated. Similar ideas were put forward earlier. For example, W. Calvin [21] introduces the concept of a "global workspace" that results from neuron interactions and integrates information from different brain regions. A. Bailey [22] examines James' theory of the stream of consciousness which includes subjective feelings and emotions. The functional model of a new-generation computer system with AI is shown in Fig. 3 with modules [23] of artificial consciousness and artificial conscience. The first attempt to create AI based on the conscience conception was made by the firm Anthropic [24]. Anthropic claims that their Claude chatbot adheres to many rules, including the principles enshrined in the Universal Declaration of Human Rights. The firm Anthropic claims that its chatbots have a "conscience".

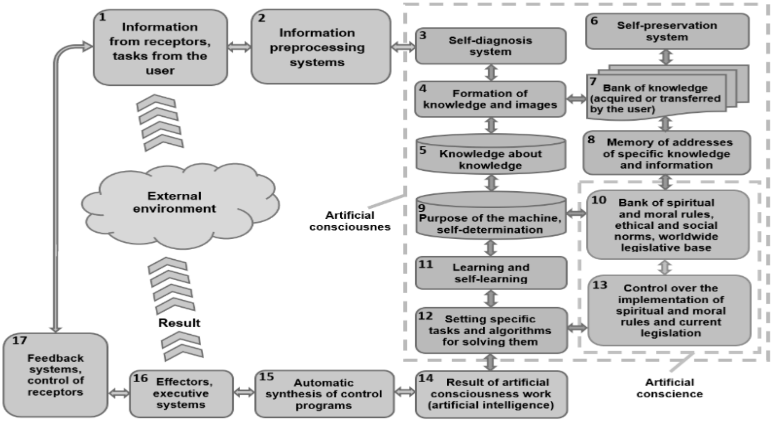


Fig.3. New-generation computer system with embedded artificial consciousness and conscience

Creating cognitive computers and robot knowledge analysis based on cognitive computing and modular neural networks [25, 26] is also a very promising research direction in AI design. Cognitive computers are intelligent processors advanced from data and information processing to autonomous knowledge learning and intelligence generation [25]. Many scientific publications are devoted to the research and design of cognitive computers. For example, the work [25] (based on deep analysis) presents a retrospective and prospective review of the odyssey toward cognitive computers empowered by transdisciplinary basic research and engineering advances. For now, a wide range of fundamental theories and innovative technologies for cognitive computers is explored, and a set of underpinning intelligent mathematics is created. For example, the architectures of cognitive computers for Cognitive Computing and Autonomous Intelligence Generation are designed as a brain-inspired cognitive engine [25]. The development of new solutions and implementation of cognitive computers in autonomous AI have good perspectives in the near future.

4. New solutions and research directions in intelligent robotics

Ukrainian scientists and policymakers pay special attention to the development and implementation of robotic systems. In intelligent robotics, research is focused on the evolution of autonomous ground vehicles (UGVs), autonomous surface (USVs), underwater vehicles (AUVs), unmanned aerial vehicles (UAVs), and the integration of these systems. The development of all these directions is propelled by the implementation of advanced AI capabilities. Machine learning is utilized for navigation, obstacle avoidance, and decision-making. There is an emerging trend towards employing multi-robot systems featuring dynamic self-organization of swarms. In underwater robotics, equipping AUVs with advanced sensors and AI algorithms assists in pipeline inspection/protection and research of marine biology, underwater environments, and landscapes. Current developments aim to improve the autonomy of AUVs, allowing them to operate for extended periods and at considerable distances in challenging underwater conditions. Aerial robotic systems, especially drones or UAVs, are used for a wide range of applications, from delivery services to combat operations. Future research in this field focuses on swarm robotics, where a group of drones collaboratively performs complex tasks. Alongside this, there is a need to develop AI algorithms that facilitate drone navigation in complex urban conditions, particularly in the absence of satellite navigation signals. An important direction for future research is the development of integrated robotic systems capable of operating in terrestrial, underwater, and aerial environments. This could potentially involve the creation of amphibious robots or systems where terrestrial, underwater, and aerial robots work in harmony. As robotic systems gain greater autonomy, accompanying ethical and legal issues must be addressed. Specifically, this concerns responsibility for the consequences in case of accidents, privacy issues related to surveillance drones, and ethical implications of autonomous weapon systems. Let us outline the future trajectories for the development of intelligent robotic systems (IRS). Direction 4.1. Powerful LLMs such as GPT-3.5 and GPT-4 developed by OpenAI [27] use machine learning to generate human-like text and have found diverse applications in IRS. Implementation of such LLMs can enable robots to understand commands given in natural language, generate human-like responses, and even engage in conversations, making them more useful and easier to use. Direction 4.2. By learning to understand a set of rules or guidelines laid out in natural language, a robot gains the ability to use these rules to make decisions in real-world situations. In this way, a prototype of an artificial conscience could be implemented, whose mechanism would allow for the avoidance of issues in communication with humans and other robots, making decisions that respect human rights and universal values. Using LLMs robots can be trained to read and understand instructions, allowing them to learn to perform new tasks without the need for explicit programming. A robot can use a GPT model for role-playing behavior, generating a series of potential actions and estimating probable outcomes of the activity, and then choosing the action most likely to achieve its goal. Direction 4.3. Additional possibilities will be provided by the development of local GPT analogs such as LLaMA, Alpaca, etc. These models can be embedded in the onboard equipment of a robot, increasing its independence from external communication networks. This direction is closely related to the development of neural networks designed for converting audio streams into text and text-to-speech. Although LLMs offer many potential advantages for the IRS, they also pose certain challenges. For example, GPT models may sometimes generate incorrect or nonsensical responses, their training requires large volumes of data and computational resources, and their use in autonomous systems raises significant concerns regarding safety, ethics, and legal regulation. These issues will require thorough investigation to ensure the beneficial use of such AI technologies without posing excessive risks.

Direction 4.4. Importantly, (a) future solutions for IRS will allow robots to learn from past experiences and adapt to new situations without human intervention; (b) improvements in sensor technology can enable robots to better understand and interact with their environment [2, 5] as robots will be able to detect and respond to changes in temperature, pressure, light, and other environmental factors. Quantum sensors (QS), which employ quantum physics to measure physical quantities with unprecedented accuracy, will significantly influence the future of IRS. QS could detect minute changes in environmental conditions, such as local fluctuations in temperature, pressure, or magnetic fields, and can provide high-precision measurements of acceleration and rotation, making them useful for robot navigation in situations where GPS is unavailable or unreliable, such as underwater or in space. QS might also enable the generation of new forms of imaging, such as ghost images, where the image is formed from light that has never interacted with the object. This technology could be implemented in robot vision systems, enabling them to see through obstacles or around corners. Furthermore, QS could play a role in secure quantum communication. This would allow robots to communicate with each other and with control systems in a way that is secure from eavesdropping. The collaborative robots concept (cobots) is expected to evolve, allowing smoother interaction between humans and robots. Robots will be able to anticipate human actions and respond accordingly. Distributed robotics, involving multiple robots working together to achieve a common goal, may see advancements leading to more complex swarm behaviors and increased task efficiency. Another subfield of robotics destined to make a place in society is soft robotics. This field involves the construction of robots using highly pliable materials, leading to the creation of more flexible and adaptable robots. Faster and more reliable communication between robots and their control systems will be facilitated by the transition to 6G - the future generation of wireless technologies.

Direction 4.5. A significant impact on the future of the IRS will be made by Augmented Reality (AR) technology [28]. It will provide a more intuitive way for humans to interact with robots, visualizing a robot's intentions, planned actions, or internal state. This will simplify the understanding and prediction of a robot's behavior for humans. AR and Virtual Reality (VR) can also be used to create realistic training scenarios for robots, which is particularly beneficial in situations where real-world training would be dangerous, expensive, or impractical. Furthermore, AR could enable remote control of robots, allowing a human operator to use an AR headset to see the same view as the robot and control the robot's actions from a safe distance. AR will enhance robots' autonomy by helping them better understand and navigate their environment. It is important to note that the realization of the outlined future achievements in the IRS will depend on various factors, including scientific breakthroughs, actual advances in the development of respective technologies, societal perception, as well as legal and ethical approaches.

5. AI perspectives in the marine industry

Implementing AI technologies has a good perspective on the marine industry, which is very important for Ukraine as a marine country. For example: (a) the AI multi-software complexes are successfully used in design processes in shipbuilding and ship-repairing; (b) intelligent polymetric sensor systems are highly efficient as information components of the integrated ships' control systems [29]; etc. Particular attention should be paid to increasing the efficiency of ship safety monitoring systems based on AI [30, 31], which can: (a) provide the seafarer with reliable and visualized factual information concerning ship loading and wind-wave impact to increase the soundness of his decisions for safe and efficient routing in heavy sea conditions, in particular, to provide navigational safety at stormy seas; (b) control of the autonomous (crewless) marine vehicle for fulfilling its mission with correction of the planned path, speed, and course in the current sea environment. Early and current research on ship safety monitoring systems focused primarily on using sensors and other hardware devices to detect hazards such as collisions, fires, capsizes, and leaks. Such systems effectively recognize potential dangers and warn crew members and other stakeholders early. However, these systems' existing hardware/software limitations could influence their effectiveness. For example, these systems were often limited by the hardware devices' processing power, which can result in delays in data analysis and decision-making. In any case, they still are DSS - Decision Support Systems [5, 31, 32], proposing to the ship Master only visualized forecasted options and limitations on routing choice (Fig. 4 and Fig. 5, where green zones mean safe parameters and red zones – dangerous parameters). The latest developments in Smart Sensors and AI-based systems for ship safety monitoring have shown promise in addressing these limitations. By leveraging intelligent digital sensors and AI technology instruments, these systems can provide real-time data analysis and predictive capabilities that can improve safety and reduce the risk of accidents. Direction 5.1. The AI-based ship safety monitoring systems involve AI algorithms and advanced digital sensors to detect and analyze potential hazards in real-time operations and optimize ship routing from human, technical, commercial, and ecological safety points of view. This AI approach would involve the integration of multiple digital and smart sensors, including liquid cargoes, green fuels and technological liquids state parameters sensors, dynamic parameters of control units, weather conditions monitoring, main engine and auxiliary systems parameters monitoring, and other devices, to provide real-time data on conditions onboard the ship into cloud databases. The data collected by these sensors would be analyzed using AI algorithms to identify potential hazards and provide early warning to crew members and other stakeholders. The AI algorithms ship used would be designed to learn over time, improving their ability to detect hazards and provide accurate predictions, and also recommendations to create the next ships' generations.

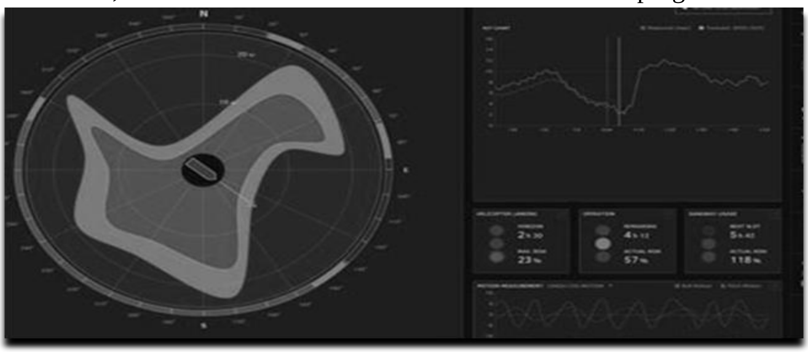


Fig. 4. Visualization of ship safety diagram: limitations by speed

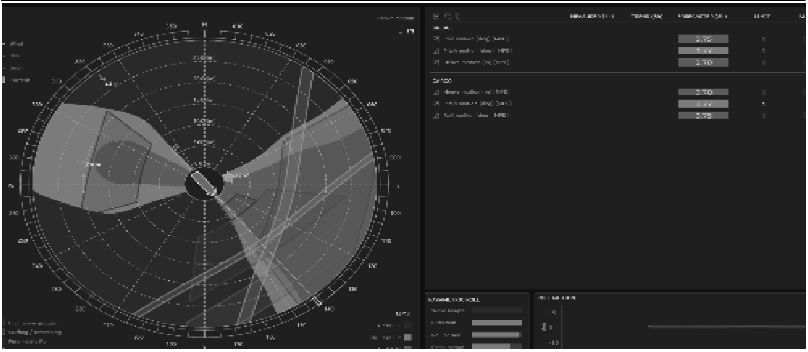


Fig. 5. Visualization of ship safety diagram: limitations by course

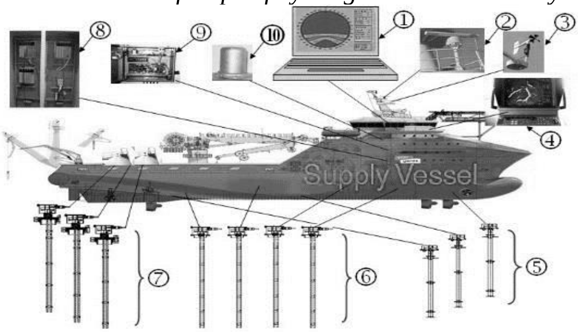


Fig. 6. Component structure of the Cyber-physical System “Supply Vessel”

The proposed AI approach's practical implementation would involve installing advanced sensors and AI algorithms onboard the ship. This approach would require the development of new intelligent hardware devices and software applications that can support real-time data analysis and predictive capabilities. For example [29], the generalized sensor-information components of the Cyber-physical System “Supply Vessel” are presented in Fig. 6. The main sensory sub-agencies of the Cyber-physical System (Fig. 6) are: operator workplace (1); a radar antenna (2); an onboard anemometer (3); the radar display and a keyboard (4); a set of sensors for ship draft monitoring (5); a set of polymetric sensors for fuel-oil, ballast water and other liquid cargo quantity and quality monitoring and control (6);a set of polymetric sensors for liquefied LPG or LNG cargo quantity and quality monitoring and control (7); switchboards of the subsystem or actuating devices and operating mechanisms control (8); a basic electronic block of the subsystem for liquid, liquefied and loose cargo monitoring and control (9); specialized electronic block with sensors for real-time monitoring of parameters of ship dynamics (10).

Potential challenges in implementing this AI approach include the need for significant investment in research and development, as well as the need for specialized training for crew members and other stakeholders. However, the potential benefits of this approach, including improved safety and reduced risk of accidents, make it a worthwhile investment for ship owners and operators.

Direction 5.2. Digital Twins (DT) would be widely used in ship design, manufacturing, operation, maintenance, modernization, repair, and, finally, their utilization. The output of DT returns to the ship as actions, recommendations, and even control. The outcomes of DT are used to: Increase safety and reduce operational costs; Design new green and digitalized ships, equipment, etc.;Training of operators and predictive maintenance; Oversight and compliance monitoring, emergency response, etc. While there are challenges to implementing this AI approach, the potential benefits make it a worthwhile investment for the shipping industry. Further research and development in this area are needed to realize the potential of AI-based ship safety monitoring systems fully.

6. Advanced AI implementation in education

LLMs ChatGPT and GPT-4 have significant potential for their use in the educational field. The capabilities of GPT-4 increased after the introduction of access to the paid version of this language model to Internet resources and the provision of the possibility of using about 800 embedded plugins, the list of which is constantly expanding. Based on the gained practical experience of working with ChatGPT and GPT-4, it is possible to formulate a set of proposals or directions regarding potential areas of application of AI platforms built on LLMs in teaching, learning, and research processes.

Direction 6.1. GPT-4 can act as a personal tutor that provides information from different areas of knowledge and can explain concepts, helping students to better understand the learning material. AI-based platforms can create a more personalized and flexible learning environment for distance learning students [33]. GPT-4 and similar technologies facilitate access to education for those living in remote areas or for individuals with special needs by converting text to speech for students with visual impairments or prov.

Direction 6.2. GPT-4 can be used 24/7 to dialogue with students, form answers to their questions, and even host seminars and discussions, allowing students to learn on their own schedule and at their own pace. AI can play the role of a personal assistant that helps navigate learning and provides advice on choosing a course, career path, or even personal development. AI should be seen as an effective means of providing emotional support to students to help them cope with stress and maintain mental health. In essence, such language models are a powerful tool for personalized learning, providing valuable experiences for the purposeful formation of knowledge, skills and abilities by adapting to the individual needs of each student. They can assess students’ current level of understanding of the learning material and provide specialized content and exercises to help them improve. In fact, we are talking about the mass creation of neuro-curators (neuro-teachers) who will answer students' questions in Telegram bots and on websites, including with voice support. The GPT-4 usage service already provides access to 5 tutoring plugins, including: (a) Tutory (affordable tutoring and on-demand education), (b) Open Lecture (offers useful points in open course lectures for targeted learning), (c) Giga Tutor (personalized AI-based tutor, which stores personalized answers to questions), (d) edX (finds courses and content from leading universities to expand the user's knowledge at any level), (e) ABCmouse (offers fun and informative learning activities for children 2-8).

Direction 6.3. AI can help educators assess assignments, provide feedback to students, and identify areas where students are struggling and need additional support. AI also can provide teachers with resources for professional development, such as suggestions for improving teaching strategies or information that will allow them to stay abreast of the latest research in their field. Using GPT-4, teachers can automate some aspects of the learning process, reducing their workload. For example, GPT-4 is capable of automatically grading assignments or creating learning content by generating learning materials such as: lecture summaries, self-test questions, interactive practice exercises, quizzes, lesson plans, examples and scenarios to illustrate complex topics and more. It is very useful to involve AI as an aid in the review process to identify potential problems in the manuscript, such as plagiarism or inconsistencies, to develop suggestions for improving the clarity and coherence of the writing.

Direction 6.4. Another area of AI implementation is the effective management of the resources of educational institutions, for example, the optimal distribution of the teaching load among teachers, the preparation of lesson schedules, the management of library resources, or the coordination of services for students. An urgent task is the formation of the necessary background for the future involvement of artificial intelligence technologies in the process of generating compromise solutions in typical situations of managerial activity. This approach will allow for generalizing the best management practices and updating conceptual approaches to their implementation.

Direction 6.5. LLMs are capable of working with different languages, allowing for the creation of multilingual learning resources and providing global access to education. AI can be a valuable tool for native and foreign language learning, providing instant feedback on grammar, pronunciation, and vocabulary. It can also facilitate the practice of speaking in a safe and non-judgmental environment, without judgment, providing real-time translation services that facilitate the learning and collaboration of students who speak different languages. In this context, it is worth noting the additional capabilities of GPT-4, which are provided, for example, by plugins Speak (a language tutor based on artificial intelligence that allows you to learn to speak anything in another language), Speechki (provides the conversion of texts into ready-made audio with a download link sound files, or to the page of the audio player).

Direction 6.6. A promising trend is the integration of LLMs with generative transformers capable of synthesizing two-dimensional and three-dimensional images, and videos. This opens a wide field for creativity and improvement of teaching and learning processes.

Direction 6.7. Using artificial intelligence as a teaching tool, teachers and faculty can train students to identify and analyze the limitations inherent in artificial intelligence, which will help develop critical thinking. As artificial intelligence becomes more common, it is important for students to understand how such systems work and the ethical considerations involved in their use. Therefore, AI should be used as a tool to teach relevant important skills.

Direction 6.8. AI can facilitate collaborative learning by coordinating group projects, creating an environment for discussion, and providing feedback on group dynamics. A very valuable asset is the ability of AI to generate dynamic simulations or scenarios that allow students to apply their knowledge in a virtual environment. This is especially useful in fields such as medicine, engineering, or any other discipline where real-world application of skills is key.

Direction 6.9. No less important direction is psychological assistance to students. AI should be seen as an effective means of providing emotional support to students to help them cope with stress and maintain mental health. This may include mindfulness exercises, stress management techniques, or even simply providing compassionate communication. An intelligent chatbot can help identify situations where a student is experiencing emotional difficulties and suggest appropriate resources or interventions. In this context, AI can play the role of a personal assistant that helps navigate learning and provides advice on choosing a course, career path, or even personal development.

Direction 6.10. LLMs are also a valuable tool to assist researchers by providing quick access to information, generating ideas for research and writing fragments of scientific reports, dissertations, or articles, summarizing large volumes of text, identifying key themes in the literature, and even suggesting areas for further research. AI can also help with data analysis and visualization and symbolic computation. In particular, the Wolfram plugin provides access to mathematical calculations, knowledge sampling, and real-time data through Wolfram Alpha and Wolfram Language services. Although the proposed directions have great potential, it is also important to constantly consider the ethical implications and potential risks associated with the use of AI in education. These include issues related to privacy, bias, and the quality of AI-generated content. It is critical to have safeguards in place to address these issues and use AI as a tool in a way that supports rather than replaces teachers. Again, while AI platforms offer exciting opportunities, it is important to approach them critically, ensuring that the use of AI enhances, rather than detracts from, the human element of education and research. It should be noted that the GPT-4 plugin engine relies on external programs, and GPT-4 with the plugin enabled can send them sensitive information or parts of the dialogue, including the user's country, preferences, etc. In this context, it is more reliable to use local large language models that function on the closed platform of the Ministry of Education and Science, an educational institution. Examples of these are LLaMA 7B/13B, Alpaca, etc. It is significant that the effectiveness of neural networks will grow over time in accordance with the improvement of AI technologies, increasing the quality and quantity of output data, and computing capabilities. This will encourage the revision of the concepts of training and their constant adaptation. The policymakers in different countries pay attention to the implementation of the ChatGPT in education processes. For example, the Chancellor of the nation’s largest school system, New York City Public Schools, David C. Banks said on 18 May 2023 [34] that ChatGPT caught NYC schools off guard and now, they are determined to embrace its potential and in New York public schools, students will be taught how to use AI. Direction 6.11. No doubt, the efficiency of training students in the AI field at the university level may be significantly increased in the framework of specialized integrated education environments [35] such as multi-university (academic) consortia and academic-industry consortia.

7. Linguistic competency of AI systems

Linguistic competency is a recognized sign of human intelligence. It results from linguistic intelligence in Gardner’s theory of multiple intelligences [36]. Besides, verbal communication using natural language and language acquisition are included in the list of competencies characterizing human-level general intelligence [37]. Given this, the ability to express thoughts, ideas, and suggestions using human language, which constitutes the linguistic competency of an artificial personality, is an important subtask in developing AI. In turn, “accurate report”, which is a standard behavioral index indicating consciousness in humans, is best realized through human language [38]. In a similar way, these ideas are applicable to AI, which imitates human intelligence and thinking. In this sense, the developed linguistic competency of an artificial personality able to report accurately on what is going on may be regarded as a criterion indicating the rise of artificial consciousness. The role and rise of linguistic competency follow from Potebnja’s theory of thought and language which emphasizes the importance of language in shaping human thought processes. According to Potebnja, language is not simply a tool for communication but is also a means of organizing and structuring our thoughts [39]. He argues that language is a form of thought and that the way we use language reflects our cognitive processes. Potebnja’s ideas have influenced the development of cognitive linguistics, which seeks to understand the relationship between language and thought [40]. In a similar way, these ideas are applicable to AI which imitates human intelligence and thinking. Human understanding of the text or a message is based on the meanings of the used words, which are presented in explanatory dictionaries. Intrinsically, the use of such vocabulary as a knowledge resource of AI simulates the human way to formulate thoughts. The meaning of each word can be decomposed into elementary senses and these can be deduced from the word definition or explanation available in the dictionary. This process can be described mathematically and correspondingly formalized to automatically build semantic fields [41], resulting in the technical possibility of developing a deep intelligent instrument able to assess and compare texts and disambiguate word senses. The linguistic module of artificial personality can acquire human-like linguistic competency in this way. Chomsky argues [42] that humans have an innate ability to acquire language, which is hard-wired into our brains. This idea has been influential in the development of NLP algorithms, which seek to replicate human language acquisition processes in machines and AI technologies. It is important that the AI mechanisms modeling human thought and language preserve the information contained in the processed texts. If the original language uses a non-Latin alphabet, some natural language processes require its Romanization. To be able to restore the initial text, the Latinization process should be based on scholarly (strict) transliteration, which provides simple-correspondent (one-to-one) or isomorphic correspondence between initial and Romanized graphemes. The Latinization rules using a mediator language inevitably refer to corresponding sounds in that language. That is, they are based on practical transcription rather than transliteration and, therefore, fail to preserve contained semantics. For example, the use of the English-oriented Romanization system for Ukrainian results in word form distortion and the appearance of false identities: Гальченко – Галченко (Halchenko), Тронко – Тронько (Tronko), Воронко – Воронько (Voronko), Банкова – Банькова (Bankova), Паньківська – Панківська (Pankivska), Польова – Полова (Polova), Лялько – Ліалко (Lialko), Ліліана – Ліляна (Liliana), Маріан – Мар’ян (Marian), медіана – медяна (mediana), Возіанов – Возянов (Vozianov), Гундеріан – Гундерян (Hunderian), Клаузіус – Клаузюс (Klausius), Пії – Пій (Pii), Лар’їн – Ларін (Larin), Левитський – Левицький (Levytskyi), Тоцька – Тотська (Totska), Чернятський – Черняцький (Cherniatskyi), etc. To avoid these inconsistencies [43], it is expedient to introduce in AI systems the national transliteration standard DSTU 9112:2021.

Direction 7.1. Research direction, which concentrated on increasing the linguistic competency of AI systems for a correct understanding of the contents in communications between humans and intelligent robots and between different kinds of robots in multi-robotic systems, is perspective and important for future AI development and implementation.

8. Conclusion

The main peculiarities of the developed AIDU Strategy, priorities in AI implementation, and prospective research directions in the AI field are focused on and discussed in detail. The result of the “Strategy for Artificial Intelligence Development in Ukraine” [39] implementation should deal with the creation of breakthrough technologies in the field of computer science and artificial intelligence as well as the creation of conscious AI-powered computers that make decisions considering ethical, moral, and legal norms.

One promising way to realize the AIDU Strategy is the study of artificial consciousness based on a synergetic approach. At the next step, future research must deal with software and hardware development, testing and implementation of proposed new-generation intelligent systems with AI based on the conscience conception. Another important direction (7.1) is the development of linguistic technologies, particularly those providing semantic text analysis that manifest the emergence of linguistic competency of an artificial personality. Besides, the authors analyzed and underlined the most important fields for AI implementation in Ukraine, as well as, developed, formalized and justified priority practical-research directions for future successful AI results and achievements, in particular, (4.1) – (4.5) in intelligent robotics, (5.1) – (5.2) in the marine industry, (6.1) – (6.11) in the education sphere. Scientific efforts must be concentrated on intensive AI research in the abovementioned directions to increase the role of Ukraine in the world as a highly technological country, strong marine country and country with high-caliber standards in education. IAIP-project on the AIDU Strategy may be the base as well as can be modified, extended and transformed into the final edition of the National Strategy for AI Development in Ukraine.

9. References

[1] N. Lidströmer, H. Ashrafian, Artificial Intelligence in Medicine, Springer, Cham, 2022. doi: https://doi.org/10.1007/978-3-030-64573-1

[2] Y. Kondratenko, et al., Machine Learning Techniques for Increasing Efficiency of the Robot's Sensor and Control Information Processing, Sensors 22(3) (2022) 1062. doi: 10.3390/s22031062

[3] C.C. Aggarwal, Artificial Intelligence. Springer International Publishing, 2021. DOI: 10.1007/978-3-030-72357-6.

[4] Y.P. Kondratenko, et al., Slip displacement sensors for intelligent robots: Solutions and models, in: Proceedings of the 2013 IEEE 7th International Conference on Intelligent Data Acquisition and Advanced Computing Systems, IDAACS 2013, vol. 2, art. no. 6663050, 2013, pp. 861-866. doi: 10.1109/IDAACS.2013.6663050

[5] R. Duro, et al. (Eds.), Advances in Intelligent Robotics and Collaborative Automation, River Publishers, Aalborg, Denmark, 2015. DOI: https://doi.org/10.13052/rp-9788793237049

[6] The Global AI Strategy Landscape, 2019. [Online] https://bit.ly/3LFvGY9.

[7] D. Castelvecchi, How will AI change mathematics?, Nature 615 (2023) 15-16.

[8] C. Stokel-Walker, R. Van Noorden, The promise and peril of generative AI, Nature 614 (Feb. 9, 2023) 214-216.

[9] S. Russell, AI weapons: Russia's war in Ukraine shows why the world must enact a ban, Nature 614 (2023) 620-623.

[10] A. Shevchenko, et al., The Ukrainian AI Strategy: Premises and Outlooks, in: Proceedings of the 12th Int. Conf. on Advanced Computer Information Technologies, ACIT 2022, Sep. 26-28, 2022, pp. 511-515. doi: 10.1109/ACIT54803.2022.9913094
[11] A.I. Shevchenko, et al., Strategy for Artificial Intelligence Development in Ukraine. Monograph. IAIP, Kyiv (2023). DOI: 10.15407/development_strategy_2023.
[12] Concept of Artificial Intelligence Development in Ukraine, 2020. [Online] https://bit.ly/3xXylXx
[13] V. Slyusar, et al., Construction of an advanced method for recognizing monitored objects by a convolutional neural network using a discrete wavelet transform, Eastern-European Journal of Enterprise Technologies 112 (2021) 65-77. doi: 10.15587/1729-4061.2021.238601.
[14] V. Slyusar, et al., Improvement of the object recognition model on aerophotos using deep conventional neural network, Eastern-European Journal of Enterprise Technologies 5/2 (113) (2021) 6 – 21. DOI: 10.15587/1729-4061.2021.243094.
[15] C.H. Lampert, et al., Attribute-based classification for zero-shot visual object categorization, Pattern Analysis and Machine Intelligence, IEEE Transactions on 36(3) (2014) 453–465. DOI: 10.1109/TPAMI.2013.140.
[16] Jia Deng, et al., ImageNet: A Large-Scale Hierarchical Image Database, in: 2009 IEEE Conf. on computer vision and pattern recognition, 2009, pp. 248–255. DOI: 10.1109/CVPR.2009.5206848.
[17] A. Bochkovskiy, et al., YOLOv4: Optimal Speed and Accuracy of Object Detection, arXiv.org, (2020).
[18] F. Peng, et al., Traffic flow statistics algorithm based on YOLOv3, in: 2021 International Conference on Communications, Information System and Computer Engineering, CISCE'2021, 2021. DOI: 10.1109/cisce52179.2021.9445932.
[19] Hugo Touvron, et al., LLaMA: Open and Efficient Foundation Language Models, arXiv:2302.13971v1, 27 Feb 2023. doi: 10.48550/arXiv.2302.13971.
[20] Zhengxuan Wu, et al., Interpretability at Scale: Identifying Causal Mechanisms in Alpaca, arXiv.org, May 15, 2023. https://arxiv.org/pdf/2305.08809v1.pdf.
[21] W.H. Calvin, The Cerebral Code: A Thinking Thought in the Mosaics of the Mind, The MIT Press (1998).
[22] A. Bailey, The Strange Attraction of Sciousness: William James on Consciousness, in: J. Scott Jordan (Ed.), Modeling Consciousness across the Disciplines, University Press of America, Lanham, MD, 1999, pp. 43–63.
[23] A.I. Shevchenko, M.S. Klymenko, Developing a Model of Artificial Conscience, in: 15th IEEE International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT'2020, vol. 1, 23-26 Sept. 2020, Lviv-Zbarazh,2020,pp. 51–54.
[24] Artificial intelligence now has a "conscience", [Online]. https://portaltele.com.ua/news/technology/u-shtuchnogo-intelektu-teper-ye-sovist.html
[25] Y. Wang, et al., The Odyssey to Next-generation Computers: Cognitive Computers (kC) Inspired by the Brain and Powered by Intelligent Mathematics, Frontiers in Computer Science (2023). DOI:10.3389/fcomp.2023.1152592
[26] V. Vychuzhanin, et al., Cognitive-impulse model for assessing complex technical systems survivability, CEUR Workshop Proceedings, Vol. 2711, 2020, pp. 571-585. https://ceur-ws.org/Vol-2711/paper44.pdf
[27] GPT-4 Technical Report by OpenAI, 27 March 2023. .
[28] Misiuk Tetyana, et al., Computer Vision Mobile System for Education Using Augmented Reality Technology, Journal of Mobile Multimedia 17/4 (2021) 555-576.

[29]Y.D. Zhukov, et al., Intelligent Polymetric Systems Industrial Applications, CEUR Workshop Proceedings, Vol. 2762, 2020, pp. 122 – 137. https://ceur-ws.org/Vol-2762/paper8.pdf
[30]Y. Kondratenko, et al., Ship Navigation in Narrowness Passes and Channels in Uncertain Conditions: Intelligent Decision Support. in: P. Shi, et al. (Eds.), Complex Systems: Spanning Control and Computational Cybernetics: Foundations, volume 414 of Studies in Systems, Decision and Control, Springer, Cham, 2022. DOI:0.1007/978-3-030-99776-2_24
[31]V. Timchenko, et al., Decision Support System for the Safety of Ship Navigation Based on Optical Color Logic Gates. CEUR Workshop Proceedings, Vol. 3347, 2022, pp. 42–52. 4th International Scientific Conference "Information Technology and Implementation", IT&I'2022, Virtual, Online, 30 Nov - 2 Dec 2022, 186836
[32]O.V. Kozlov, Optimal Selection of Membership Functions Types for Fuzzy Control and Decision Making Systems, in: Proceedings of the 2nd International Workshop on Intelligent Information Technologies & Systems of Information Security, Khmelnytskyi, Ukraine, IntelITSIS 2021, CEUR-WS, Vol-2853, 2021, pp. 238-247.
[33]V. Shebanin, et al., Development of the Mathematical Model of the Informational Resource of a Distance Learning System, in: O. Chertov, et al., (Eds.), Recent Developments in Data Science and Intelligent Analysis of Information. Proceedings of the XVIII International Conference on Data Science and Intelligent Analysis of Information, June 4–7, 2018, Kyiv, Ukraine. ICDSIAI 2018, Advances in Intelligent Systems and Computing, Vol. 836, Springer International Publishing, pp. 199-205, 2019. DOI: 10.1007/978-3-319-97885-7_20
[34]ChatGPT caught NYC schools off guard. Now, we're determined to embrace its potential. 2023. [Online] https://ny.chalkbeat.org/2023/5/18/23727942/chatgpt-nyc-schools-david-banks
[35]G. Kondratenko, et al., Fuzzy Decision Making System for Model-Oriented Academia/Industry Cooperation: University Preferences, in: C. Berger-Vachon, et al. (Eds.) Complex Systems: Solutions and Challenges in Economics, Management and Engineering, volume 125 of Studies in Systems, Decision and Control, Springer, Berlin, Heidelberg, 2018, pp. 109-124. DOI: https://doi.org/10.1007/978-3-319-69989-9_7
[36]H. Gardner, Intelligence reframed: Multiple intelligences for the 21st century. New York, Basic Books, 1999.
[37]S. Adams, et al., Mapping the landscape of human-level artificial general intelligence. In: AI Magazine 33(1), 25–42, 2012.
[38]A.K. Seth, et al., Criteria for consciousness in humans and other mammals, Consciousness and Cognition 14 (2004) 119-139.
[39]A. A. Potebnja, Myslj i jazyk [Thought and language]. 2nd edition, Kharkiv, Adolf Darre Tipografija, 1892.
[40]D. Geeraerts, Cognitive linguistics: Basic readings. Walter de Gruyter, 2010.
[41]M. Vakulenko, From Semantic Metrics to Semantic Fields, in: Proceedings of the 2021 IEEE 16th International Conference on Computer Science and Information Technologies, CSIT'2021, 22–25 September 2021, Lviv, Ukraine, 2021, pp. 44–47. DOI: 10.1109/CSIT52700.2021.9648675.
[42]N. Chomsky, Language and Mind, 3rd edition, Cambridge University Press, Cambridge, 2006.
[43]M. Vakulenko, Normalization of Ukrainian letters, numerals, and measures for natural language processing. In: Digital Scholarship in the Humanities. Published online: 29 December 2022, (2022). DOI: 10.1093/llc/fqac090.

УДК 004.8

ОСНОВНІ НАПРЯМКИ РЕАЛІЗАЦІЇ СТРАТЕГІЇ ШТУЧНОГО ІНТЕЛЕКТУ В УКРАЇНІ

Член-кор. НАНУ, Dr.Sci. Шевченко А.[1], Dr.Sci. Кондратенко Ю.[1,2], Dr.Sci. Слюсар В.[1], Dr.Sci. Жуков Ю.[3], Ph.D. Кондратенко Г.[2], Ph.D. Вакуленко М.[1]

[1]Інститут проблем штучного інтелекту Міністерства освіти і науки та Національної Академії наук України, Україна,
E-mail: rector_iai@ukr.net, ypk@ipai.net.ua, swadim@ukr.net, maxvakul@gmail.com
[2]Чорноморський національний університет імені Петра Могили, Україна,
E-mail: y_kondrat2002@yahoo.com, halyna.kondratenko@chmnu.edu.ua
[3]C-Job Nikolayev, Україна,
E-mail:y.zhukov@c-job.com.ua

Анотація. *Даний розділ присвячена аналізу конкретних напрямків та особливостей Стратегії розвитку штучного інтелекту (ШІ) в Україні. Основними складовими роботи є аналіз поточного стану з обґрунтування, розробки та державного затвердження Національної стратегії ШІ в Україні; ключові елементи та основні пріоритетні напрямки впровадження ШІ згідно з ІПШІ-проектом «Стратегія розвитку ШІ в Україні»; пропозиції щодо розвитку ШІ в короткостроковій та довгостроковій перспективі та особливості впровадження ШІ в Україні в сучасний воєнний час. Особлива увага приділяється таким напрямкам досліджень ШІ, як (а) проектування систем ШІ на основі когнітивних концепцій і концепцій свідомості; (б) нові рішення в інтелектуальних робототехнічних системах для наземного, підводного та повітряного застосування; (в) перспективи штучного інтелекту в морській галузі; (г) перспективне впровадження ШІ в освіту; (д) лінгвістична компетентність систем ШІ. Отримані авторами результати можуть бути використані для розробки стратегічних кроків і планів у дослідженні ШІ та впровадженні процесів прийняття рішень на державному рівні.*

Ключові слова: *Стратегія, штучний інтелект, розробка, впровадження, Україна, пріоритети, особливості, аналіз, ІПШІ-проект*

UDC 519.85

APPROACH FOR DETERMINING THE INFLUENCE OF UAV SPATIAL MOTION PARAMETERS ON THE CHARACTERISTICS

Lysenko O.[a], Dr.Sci., Tachinina O.[b], Dr.Sci., Kirchu P.[c], Ph.D., Ponomarenko S.[a], Ph.D. Kutiepov V.[b], Sushyn I.[a]

[a] *National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine*
[b] *National Aviation University, Ukraine*
[c] *«UTAS»Company», Ukraine*
ORCID: 0000-0002-7276-9279. 0000-0001-7081-0576, 0000-0001-8437-4402, 0000-0001-5512-3778, 0000-0002-1055-9698, 0000-0003-4866-4351
E-mail: tachinina5@gmail.com, lysenko.a.i.1952@gmail.com, guydasg@ukr.net, fkirchu@gmail.com, novikov1967@ukr.net, rubin268@ukr.net

Abstract. *The key element of the latest broadband wireless access systems operating in the conditions of a metropolis, mountainous or cross-country terrain, as well as in the conditions of combat operations and emergency situations is a high-altitude aerial platform with special telecommunication equipment located on it. The use of an aircraft-type small-sized unmanned aerial vehicles (UAVs) for such systems allows to achieve the following advantages: relatively low cost of implementation, small size, a short period of time for the deployment of communication systems in a given area. The section examines the fields of implementation and the principles of structural and functional construction of telecommunication systems based on high-altitude aerial platforms. An analysis of the principles of structural and functional construction of telecommunication systems based on aerial platforms was carried out, which showed the main perspective areas of application of wireless episodic networks, where one of the structural units is aerial platforms on which repeater nodes are placed. This makes it possible to provide a given quality of service (QoS) to subscribers by maintaining connectivity between network nodes. The functioning of such complex networks is not possible without an effective control system, which includes the UAV network control subsystem and the flight control subsystem. Conducted researches in the field of development of UAV control systems made it possible to determine the basic advantages and disadvantages of the latest approaches to the synthesis of flight control laws, and the conditions for their application. The selection of the optimal adaptive control system of the aerial platform made it possible to choose a quadratic functional as an optimization criterion. This will make it possible to ensure the minimization of energy consumption while maintaining the necessary quality indicators of the functioning of the telecommunications system, in the conditions of changing the parameters of the control object.*

Keywords: *UAV, aerial platform, telecommunication network**,** wireless episodic network*

INFORMATION PROCESSING IN CONTROL AND DECISION-MAKING SYSTEMS. PROBLEMS AND SOLUTIONS

1. Introduction

In conditions of high probability of local (regional) armed conflicts, the developed countries of the world pay special attention to the improvement of communication systems of the tactical link of management [1]. Achieving information advantage is presented as an objective necessity for the successful course of the battle (operation). The main features of modern warfare can be emphasized: deployment of battle formations on a wide front and great depth, rapidity, high maneuverability, the use of high-precision weapons, a global intelligence system, new ways of conducting military operations, deep maneuver raids, etc. The main features of the process of managing troops, combat systems and weapons are:

– increased mobility of units;

– high dynamics of movements of troop groups as a whole;

– decentralized deployment of troops in territories divided by enemy forces;

– integration of communication, navigation, intelligence and automation systems, etc.;

– a single information space for all its participants;

– focus on direct participants in military operations (automation of the battalion – company – platoon – individual soldier levels);

– decentralization of network resource management processes.

It is known that modern principles of communication organization and technical equipment of communication units of Ukrainian Armed Forces do not allow to fully satisfy the needs for commanding of troops in the conditions of modern combat [2] The main disadvantages of the existing communication system of the tactical control link are: low mobility of communication nodes of control points; non-fulfillment of requirements for productivity, reliability, security, providing of radio communication between mobile subscribers; non-fulfillment of the probability-time characteristics of information exchange; low automation of the processes of establishing, maintaining and supporting radio communication; moral and physical obsolescence of radio communication equipment, etc. This means, that the development of new approaches to the creation of a tactical-level communication system that meets today's requirements is an actual scientific problem. Therefore, a new architecture of the mobile component of the communication networks of the Ukrainian Armed Forces is proposed, which takes into account: the experience of the development of tactical communication systems of the United States; modern requirements of the troop management process and corresponding requirements for tactical-level communication systems; the modern level of development of wireless telecommunication technologies for civilian purposes.

2. Analysis of previous studies

Recently, the USA has been intensively developing a single multifunctional information and control system that integrates the functions of managing troops, weapons, intelligence, electronic warfare, as well as communication, navigation, orientation and recognition - C4ISR (Figure 1) (Command, Control, Communications, Computers, Intelligence, Surveillance & Reconnaissance). The architecture of C4ISR is presented in 1.1. Since 1999, the term network-centric concept of warfare has been introduced, which defined the large-scale use of computers, high-speed channels and network software on the battlefield.

Classic analog radio communication is actually no longer used in the corps, brigade and battalion units of military command. It was replaced by wireless information networks that allow receiving not only formalized reports on discovered and destroyed targets, losses, ammunition and fuel consumption, but also video images from the scene of combat area, information from reconnaissance unmanned aerial vehicles, radio electronic surveillance aircraft and ground surveillance targets.

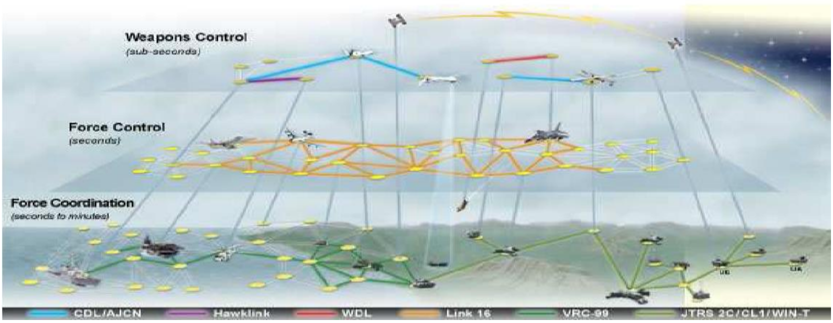


Fig. 1. C4ISR architecture

In the US armed forces, a phased implementation of the battlefield information network (WIN-T - Warfighter Information Network Tactical) program is proposed, which involves the reorganization of divisions into "computerized" ones. The purpose of the reorganization is to reduce the combat and numerical composition of the division with a simultaneous increase in its combat effectiveness due to increased mobility, achieving an absolute advantage in information provision and intelligence capabilities. A characteristic feature of the "computerized" division is equipping it with promising models of weapons, military equipment, automatic control system of units, intelligence, air defense and rear, integrated into one system.

2.1. Presentation of main material

In our time the rapid growth of technologies in the field of telecommunications has led, on the one hand, to the anticipatory development of commercial networks, on the other hand, the need to reduce budget allocations for defense requires the introduction of civil technologies into the military sphere. However, the analysis of existing standards (IS–54, IS–41, IS–95A) and protocols (IPv.4, GSM MPT 1327, TETRA, etc.) of cellular and trunked radio system, wireless local shows that their direct application in tactical systems of military communication is impossible. The main difference of tactical level networks is unreliability, temporality of channels, their low survivability; at the same time, technologies, protocols, quality of service (QoS) of the Internet are designed for the intended infrastructure at the physical level. Commercial protocols are optimized for stationary infrastructure (base stations are stationary) and cannot perform the functions of addressing, routing, service transfer from one zone to another in networks with a dynamic topology. They implement centralized control algorithms and, in addition, bring a significant service load to the network (Table 1). Therefore, in promising tactical networks, it is necessary to use commercial standards and carry out scientific developments taking into account the unique characteristics of military infrastructure. The features of modern tactical communication networks include: – dynamic topology (network nodes are mobile, subject to destruction and failures; radio communication channels are unstable, have limited communication range and bandwidth due to the influence of the enemy's radio-electronic countermeasures, mutual interference, conditions of radio waves propagation, etc.); – limited power and transmission time of subscriber equipped with radio terminals with rechargeable batteries; – significant size of networks (hundreds or thousands of elements); - heterogeneity of network elements: by mobility (tank, soldier, helicopter, aircraft), by level of productivity (mobile base station, mobile subscriber).

Table 1

Variants for building the architecture of tactical link networks

Characteristic	Cellular networks	Trunked networks	Mobile radio networks
Architecture	Fixed cellular: fixed service areas, stationary base stations		There is no fixed infrastructure, each node is a relay (router) of messages
	Usage of public stationary network	Using a fixed network to connect base stations	
Type of topology	Static (base stations are static)		Random, highly dynamic, topology adaptation to operating conditions
Deployment time	Very long	Significant	Rapid deployment,
	A network design (planning) stage is necessary	A network planning stage is necessary	self-organizing of network, easy expansion
Control type	Centralized, the presence of a separate (dedicated) control network		Decentralized, does not have a dedicated control network
Mobility	Only subscribers within the coverage areas of stationary base stations are mobile		All network elements are mobile
Survivability	Very low		Very high
Intelligence security	Low		High
Data transfer speed	Low speed - GPRS (one - ten Kb/s)		High, 1.. .54 MB/s
Field of application	Rear units, anti-terrorist operations		Battlefield, emergencies

At the same time, the main technical requirements for the next generation of communication systems are: – integration of all types of traffic (voice, data, video, video conference); – full mobility of all subscribers and network elements; - ensuring the specified quality of subscriber service (QoS) in significant geographical areas under conditions of use of both conventional and nuclear, biological and chemical weapons; - guaranteed confidentiality of all types of information; - minimal human participation in planning and communication. Work [3] shows the analysis of possible options for building the architecture of tactical link networks demonstrated the advantages of using mobile Ad-hoc Networks (MANET) in comparison with cellular or trunking networks – table 1. Mobile ad hoc network (MANET) is a dynamic architecture of building networks which is self-organizing and does not contain base stations and fixed information transmission routes. A network node is a terminal (portable computer, personal secretary, sensor device, robot, etc., equipped with a radio modem) which utilizes the functions of a host and a router. In these networks, the topology is random, all its elements can be mobile, the principle of organization of information transmission is the switching of messages (packets), control type is decentralized. The main characteristics of the existing channel-level protocols of wireless networks are given in table 2. The analysis of the table allows us to conclude, that in conditions of high mobility of nodes, which is characteristic of most BEM, the IEEE 802.11 and IEEE 802.16 protocols are preferable, although they are far from perfect, but after certain modification they can be used in tactical radio networks. There are already certain commercial technologies (standards) of the channel and physical levels of the OSI model, which can be the basis for creating a mobile component of the tactical communication networks of the Ukrainian Armed Forces. The mobile component is designed to ensure information exchange in the interests of all troops operating in the tactical zone, regardless of their subordination and the tasks they perform. It is assumed that its architecture will be heterogeneous, hierarchical, consisting of three main levels (Figure 2): 1st - mobile radio networks of the lower level of management; 2nd - networks of mobile base stations (MBS), which will form a support network; 3rd is an aerial network that can be implemented on unmanned aerial vehicles [2].

Table 2

Characteristics of the existing protocols of the channel level of wireless networks

Main characteristics	Channel level protocol			
	IEEE 802.11	IEEE 802.16	HiperLAN 2	Bluetooth
Frequency (GHz)	2.4 / 5.1	2 - 66	5.1	2.4
Distance (m)	From 500 / to 100	100 - 20000	From 250	10...100
Transmission speed in the channel Mb/s	1 / 2 / 11 / 54	120	54	0.7...1
Channel access method	DFWMAC (CSMA/CA)	OFDMA/TDMA /Polling / TDD TDD		
Control type, network organization	Decentralized, all nodes are of the same level	Zonal (network clustering), centralized management of resources in each zone (piconet)		
Node mobility	Easily implemented at the network level	Complicated, causes the need to rebuild network zones		
Network bandwidth	Limited by mutual interference	Determined by the network configuration, network reconfiguration time is significant		

Sensor networks (telemetry networks) can form an additional zero level. Creation of each level involves improvement of quality indicators of functioning of the entire communication system. Each level of the mobile component uses its own frequency subband. The advantages of mobile radio networks are obvious: no planning stage (possibility of self-organization), fast deployment, decentralized management (very high survivability), work in motion of all network elements, etc. For communication between geographically separated groups of troops (network zones) or to increase the reliability of communication between the MBS and the productivity of the mobile component, an upper level is created - an air trunk network that can be implemented on unmanned aerial vehicles (airplanes, airships). Proposals for the creation of a UAV network in the interests of all of Ukraine can be found in [2]. Each UAV is equipped with two types of radio equipment using directional antennas: 1st – for communication with MBS or dedicated users; 2nd – for exchanging information with a nearby UAV.

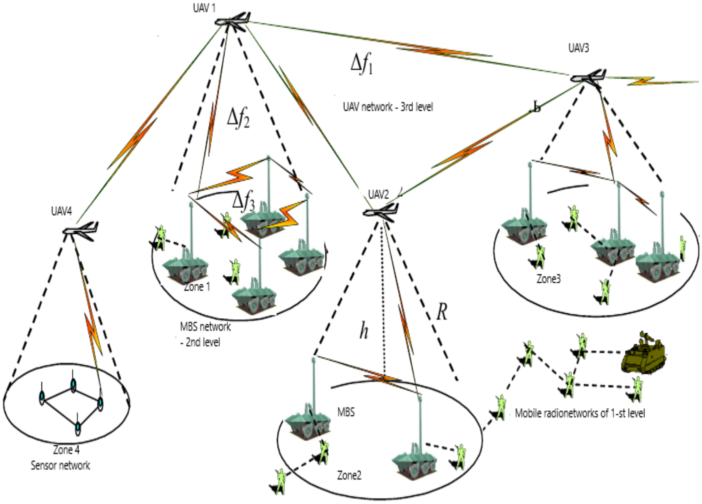


Fig. 2. An example of the architecture of the military communications network mobile component using a UAV network

UAVs are integrated into a network of aerial nodes for switching messages (packets) with the implementation of routing functions: collection (sending) of route information, its storage, calculation of routes, transmission of packets along two types of routes. The first type of route provides relaying of traffic within its zone, the second type - between different (m-n) zones. The advantages of using a UAV network are following:

1. Connectivity between geographically separated groups of troops (zones of the mobile component) is ensured [2].

2. The reliability of communication between MBS within the same zone increases due to the appearance of alternative independent transmission routes.

3. The performance of the network is increased due to: use of radio channels between UAVs with greater bandwidth compared to the MBS-MBS radio channel, efficiency of managing the mobile component (the volume of transmitted service information is reduced and the time of its collection is reduced [2]), reducing the length of information transmission

routes by several times, etc.

4. The specified quality of subscriber service (QoS) is provided due to the use of deterministic multiple access protocols [2].

5. Remote collection of intelligence information or its removal from sensors of sensor networks is provided.

2.2. Development of aerial platforms based on UAVs for telecommunication systems

Nowadays, more and more countries of the world are paying attention to the development of UAVs that can be used in various spheres of human activity. One of these areas is the use of UAVs as an aerial platform for placing repeaters of wireless episodic communication networks that are used in households and on the battlefield. Aerial platforms for telecommunication systems are aircraft lighter or heavier than air, which can move in the air space for a long enough time, carry cargo with telecommunication equipment, ensure its operability and perform, if possible, orientation of antenna systems on the Earth's surface. Depending on the working altitude, aerial platforms can be divided into three types: low-rise (up to 7..8 km), medium-rise (up to 8..12 km) and high-rise, or high-altitude (stratospheric) [4]. Usage of economic UAVs became an important stage in the creation of aerial platforms. Thus, as an air platform, General Atomic Aeronautical Systems Incorporated offers the RQ-1 Predator reconnaissance aircraft with the following characteristics: wingspan of 14.8 m; length 8.22 m; height 2.1 m; speed 200 km/h; operating altitude 7.62 km; range 742 km; time in the air is 40 hours; empty weight 500 kg; payload mass 204 kg. The UAV has a television and infrared camera, radar, communication and guidance equipment. Aircraft is controlled from Earth by an operator through a special station in the *Ku* range.

General Atomics Altus UAV was created to test new technologies for a long-term operation in the air of an unmanned aircraft (from 12 to 72 hours) with a launch mass of 940 kg and a payload of 150 kg. Aircraft is equipped with a Rotax 912-2T engine with a 2.5 m diameter propeller. American company Aurora Flight Sciences Cor. for long-term flights in the stratosphere offers the Perseus B UAV, which has the following specifications: wingspan of 21.8 m; take-off weight 1100 kg; payload mass

50..300 kg; altitude 20 km; flight range 1853 km; flight duration 8..24 hours; the power delivered to the payload is 1.2 kW. Also, the use of solar-powered UAVs as aerial platforms is a promising direction. One of them is the unique autonomous aircraft Helios. Such aircraft with a wingspan of 76 m accommodates 5 nacelles, which houses the landing gear, all electronics and electrical engineering, as well as 14 engines with a power of 2 hp each and weighing 5 kg each. A special regenerative module can be placed in the nacelle to provide energy to the aircraft at night. The total weight of the aircraft does not exceed 600 kg. In August 2001, Helios made a four-day non-stop flight and reached an altitude of 29.5 km [4]. It should be noted that the success achieved in the field of stratospheric autonomous drones allowed the company Aero Vironment Inc. create a new Sky Tower company specifically for the deployment of telecommunication systems based on aerial platforms. As an aerial platform for telecommunication systems, the Ukrainian UAV Phaeton, which is a development of the Zhukovsky National Aerospace University (Kharkov) [4]. The device is designed for long-term non-stop flight at a speed of 150 km/h at altitudes of up to 25 km. With a weight of 150 kg, length of 6.4 m, height of 1.9 m and wingspan of 15.8 m, the device can carry a payload from 15 to 60 kg. Own electricity is used mainly to power two electric motors of 1.5 kW each. Electricity on the board of the aircraft is produced by solar batteries (1.6..2 kW), which are placed on the upper surface of the wings and fuselage. The specific power of film photocells (on the wing consoles) is 150 W/m2, and of rigid photocells (on the fuselage) - 200 W/m2. At night, the operation of the aircraft is supported by a battery with a capacity of 16 kWh.

2.3. Features of the UAV as a control object, which is considered in this work

If we describe the classification of adoption support systems (DSS), then it should be emphasized that there is no single classification for them. Different authors have different approaches to the issue of DSS classification depending on the characteristics of the algorithm of work and the goals for which they are intended. For the structural-parametric synthesis of the optimal adaptive control system, the Ukrainian UAV Albatros of the Kharkiv Design Bureau *"ВЗЛЕТ"* is used in this work, which is by weight is a lightweight (13.5 kg) with a wingspan of 2.89 m and a wing area of 0.55 m2. Ensuring the relative cheapness of the complex requires the use of cheap sensors (fiber-optic and micromechanical gyroscopes and micromechanical accelerometers). This aircraft has thrust control, which allows you to control the speed of the aircraft with a specified accuracy, this is necessary to support the functioning of the telecommunications wireless network. The air speed of the UAV can vary from 54 to 180 km/h, which causes a change in the parameters of the aircraft, since the state matrix A and the control matrix B of the object include the coefficient $\rho V^2/2$, where V - is the true air speed. The change in UAV parameters, which is caused by the change in speed pressure and fuel combustion, determines the use of an adaptive control system to ensure constant indicators of the quality of control of the UAV. In figure 3 shows the structure of the navigation and piloting complex, which is used to control the above-mentioned UAV [5].

The on-board UAV navigation and control complex includes four main elements: 1. Integrated navigation system; 2. Satellite navigation system receiver (SNS); 3. Autopilot module; 4. Remote control receiver (RC)

The integrated navigation system measures orientation angles (course, pitch, roll), longitude, latitude, altitude, speed, angular velocities, linear accelerations, barometric altitude, vertical speed. The autopilot module issues control commands in the form of PWM (pulse width modulated) signals, in accordance with control laws. It should be noted that when synthesizing UAV control laws, both stochastic components are necessary, which are caused by the action of, for example, turbulent wind, and deterministic disturbances. Traditional or non-adaptive control systems include systems built using a priori information about the object, sufficient to achieve the control goal, regardless of the implemented control principle, the presence of feedback, the randomness or determinism of disturbances, and the use of computing tools. The development of the theory of automatic control since the end of the 50s is largely connected with the works of R. Kalman and R. Byusa on optimal linear filtering, as well as A. M. Letov and R. Kalman on the synthesis of linear dynamic systems, optimal according to the quadratic quality criterion.

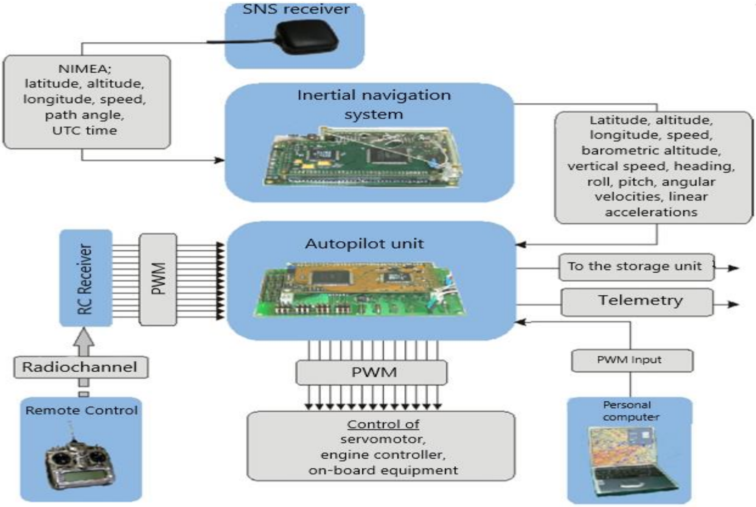


Fig. 3. On-board UAV navigation and control complex

2.4. Evolution of UAV control systems

These works formed the theoretical basis for the wide application of the theory in various fields of science and technology and made it possible to solve fundamentally new theoretical and applied problems. At the same time, practice of applying theory of optimal systems in solving specific technical problems has shown that optimal systems synthesized according to the quadratic quality criterion are sensitive to the parameters of the model of a real object and the characteristics of input signals, that is, they are not rough, and sometimes lose not only optimality, but also efficiency in those cases when a priori information about the object and the environment is not known precisely, but only with some credibility, which is specified by the membership intervals (uncertainty classes). Tasks of the synthesis of the regulator and the state estimator, taking into account the uncertainty in the object models and the characteristics of the input signals, are among the central ones in modern control theory. Their importance is primarily due to the fact that in any engineering task of designing control systems there is uncertainty (or error) in the model of the object (the mathematical model of the object, obtained on the basis of theory or as a result of identification, differs from the real technical system) and in the knowledge of the class of input disturbances. The beginning for making classical theory of robust control, extended to multidimensional systems, was the article by Zeims [6], in which a new optimality criterion was proposed based on H^{∞} - norm of the instantaneous transient function of a closed system. The use of norms as a criterion of optimality in the synthesis of multidimensional systems is based on the fact that H^{∞} norm can serve as a measure of strengthening the system. H^{∞} - the norm of the step response is the output energy of the system when a signal with unit energy is applied to the input of the system. If the output is an error, and the input is a disturbance, then by minimizing H^{∞} - norm of the step response, we minimize the error energy for the worst case of the input disturbance. New solutions to the problems of synthesis of control systems, which were reduced to the optimization of the H^{∞} norm and took into account uncertainties in the system, about which information is minimal, received their first solutions in the mid-1980s based on several approaches at once. However, sometimes synthesis procedures led to curious cases. Thus, for a second-order system, the optimal H^{∞} - regulator had the tenth order [6]. Many works published after 1984 developed the so-called "1984 approach" proposed by J. Doyle, in which, based on the theory of Glover's genkiel approximation, a procedure in the space of states for solving the problem H^{∞} - optimization for the case of finite-dimensional linear systems is given. Based on the results of the work [7], a new concept of the approach to the solution H^{∞} - optimization was formulated in 1989, named "2- Riccati approach". The point of the approach is that the optimal problem is replaced by a suboptimal one. The "2-Riccati approach" method combines the classical theory of automatic control and the state-space method, namely; the formulation of the problem was performed in the frequency domain, and its solution was performed using the state-space method. In addition, this method allows developers to set the desired characteristics of the quality and robustness of the closed system during the design process. Within this approach, the procedure of synthesis H^{∞} - suboptimal control was similar to the procedure of synthesis H^{2} -

optimal control. Applying this approach, it was possible to formulate the principle of separation in H^{∞} control theory, analogous to the well-known principle of separation in LQG theory. It was proved that, under certain conditions, H^{2} control theory is an extreme case of the H^{∞} theory. The procedure for finding suboptimal regulators was greatly simplified [8]. It was found that the degree of the regulator for an object of order n does not exceed n [8]. An alternative to robust suboptimal regulators can be adaptive automatic control systems. The emergence of the theory of adaptive systems is dates back to the second half of the 1950s, although separate adaptive systems and theoretical developments devoted to them appeared much earlier. All control systems built using a priori information sufficient enough to achieve the control goal are non-adaptive, or traditional, control systems, regardless of the implemented control principle, the presence of feedback, the randomness or determinism of disturbances, the use of computing tools, etc. If the amount of existing a priori information about the properties of the object cannot ensure the achievement of the formulated control goal, then it should be about adaptive control systems. Thus, only those control systems that are designed to function in conditions of a priori uncertainty and which in the process of functioning automatically adapt to unpredictable changes in the properties of the control object and the environment are classified as adaptive. Thus, the dynamic characteristics of aircraft strongly depend on the flight mode, technological differences, and the state of the atmosphere. In such conditions, it is often impossible to apply traditional methods, or they give poor results. Adaptive systems can be divided into two major classes: self-organizing and self-tuning. According to the level of formalization of a priori uncertainty, known adaptation approaches are divided into:

– parametric adaptation, in which a priori uncertainty is caused by insufficient knowledge of the parameters (coefficients) of the control object;

– non-parametric adaptation, in which the a priori uncertainty is not directly related to any parameters.

In both cases, the uncertainty is reduced based on successive observations of the input and output signals in the control process.

Parametric adaptation problems are closer in formulation to flight control problems, since the processes that need to be controlled have been studied to the extent that they allow us to obtain the structure of the equations of motion. According to the organization of the adaptation process, the used methods are divided into: search, which are characterized by the processes of iterative movement until the required quality of control is achieved; non-search, based on the use of some necessary conditions of the desired management quality. In systems of search adaptation, deterministic or random test signals are formed by special devices, or conditions are created for excitation in the control object of undamped oscillations, which are used as search signals. The simplest search systems are the majority of extreme systems, in which the lack of a priori information is filled at the expense of current information obtained in the form of the object's reaction to artificially introduced search (trial, test) signals. The presence of trial movements is the main disadvantage of search adaptation, as they are not always permissible under the conditions of the control object. This also applies to automatic control systems of aircraft [6]. Non-search satnavs explicitly or implicitly contain a model with the desired dynamic characteristics. The task of the adaptation algorithm is to adjust the controller coefficients in such a way as to reduce the difference between the control object and the model to zero. Such control is called direct adaptive control, and systems are called model reference adaptive systems. In the case of indirect adaptive control, the object is firstly identified, and then the corresponding coefficients of the regulator (self-turning regulator) are determined. With direct adaptive control, the adaptation circuits work in a closed loop. This allows to practice changes to the parameters of the object and the regulator. However, each self-tuning loop increases in the system order by at least one, and at the same time significantly affects the overall dynamics of the system. In the case of indirect adaptive control, the self-tuning loops work in an open loop. However, all identification errors, losses of object and controller parameters significantly affect control accuracy.

For the purposes of organizing adaptation, the following can be distinguished:

– systems with special properties, as a result of which operation the control process acquires some mandatory properties, which may include stability, sensitivity to any disturbances or errors of a priori information, the given location of the roots of the characteristic equation, etc.;

– optimal systems, ensuring the minimization of some functions that reflect the quality of the controlled movement.

Among the modern approaches to the construction of adaptive UAV flight control systems, there are many different theoretical premises and technical methods. The analysis of publications allows us to draw some conclusions about the current state and researched ways of building adaptive aircraft flight control systems.

1. The concept of building adaptive flight control systems based on a combination of identification processes and proper control is dominant.

2. In all known studies, linearized mathematical models of isolated longitudinal and lateral movements are accepted as models of the controlled process.

3. Estimation of the state of the controlled process is performed on the basis of the Kalman filter (or the issue of state estimation is not discussed - it can be assumed that in these cases the accuracy of measuring the variables of the aircraft movement is considered acceptable for solving the problem).

4. The most common identification of the parameters of the selected aircraft movement model using algorithms that implement least squares method (LSM). At the same time, such methods as the generalized least squares method are also being investigated; maximum likelihood estimation; the method based on Kalman filtering; methods that use limit cycles; method of tracking model, etc.

5. Almost all flight variables traditionally measured on aircraft are included in the number of measured signals, with a tendency to exclude signals from aerometric sensors (when measuring, such sensors use an oncoming air flow).

6. The requirements for the desired movement of the aircraft are formulated either by assigning of reference models with fixed or changing parameters, depending on the speed pressure, or by the task of quadratic functionals with the appropriate selection of weighting coefficients.

7. A variety of approaches is observed in the formation of laws for controlling the movement of aircraft. Among the ideas that influence the choice of control laws and their adaptation schemes are:

– the use of analytical connections of the real process and its reference (desired) models, which leads to "direct" consideration of the desired properties of the setting of the corresponding feedback;

– analytical design of regulators, allows to determine for linear objects linear feedbacks that are optimal in the sense of quadratic criteria;

– selection of reference (basic) settings of feedback coefficients, which satisfy the “weakened” requirements for controlled flight, but in a sufficiently wide range of flight modes (with the organization of adaptation in relatively small surroundings of these settings);

– formation and use of flight control logic schemes based on the processing of received (through direct measurement or identification of information) about flight modes.

The idea of combining different principles of automatic adjustment of flight control systems deserves attention. Thus, combining settings for environmental parameters with parametric adaptation [6] allows combining the advantages of these approaches while eliminating the disadvantages inherent in each of them individually. Software tuning based on environmental parameters ensures high speed of adaptation of the flight control system to changing conditions and simplifies emerging stability problems, and self-tuning ensures high control accuracy.

2.5. Justification of choice of an adaptive system of automatic control of the aerial platform

Based on the conducted analysis [6], it can be concluded that H^2 - the optimal regulator in the feedback loop leads to unsatisfactory functioning of a closed automatic control system in the event that strongly colored random noise enters the input of such a system. At the same time, H^∞ - optimal regulator shows conservatism (excess reinsurance) if the input disturbance is still white or slightly colored noise. That means, H^2 - regulator works satisfactorily only with an input disturbance in the form of white noise, and H^∞ - suboptimal regulator – with a deterministic signal.

In [6], the characteristics of control systems were analyzed using the following principles: control with feedback and adjustable coefficients; adaptive control with reference model using Lyapunov functions; control using neural networks. Stability, speed of convergence, operation in noisy conditions, required amount of memory, etc. were compared. The main results obtained during the comparison are given in the table 3.

Table 3

Overview of characteristics of control methods

Criterion	Control with feedback and adjustable coefficients	Adaptive control with reference model using Lyapunov functions	Control using neural networks
Stability of feedback	Worst	Best	Medium
Speed of convergence	Best	Medium	Best
Real-time operation	Medium	Medium	Best
The complexity of the control program	Worst	Medium	Medium
Tracking error	Medium	Best	Medium
Suppression of interference	Best	Medium	Medium
Robustness of model discretization	Worst	Medium	Best

Neural network and adaptive control methods have no system linearity constraints, they are effective in noisy conditions, and provide real-time control after training, but neural network controllers are currently much more expensive than conventional DCMs that can be used to implement an adaptive algorithm. Therefore, the use of searchless systems with parametric adaptation is more promising for UAV flight control. In [9], it was shown that the simplest forms of adaptation, which became widespread in serial SAC [10], and competing forms of stabilization of dynamic properties, which played and still play a prominent role in SACs of some classes [10], do not solve the above-mentioned modern problems of aircraft control. To some extent, this is illustrated in figure 4. A number of principles of adaptation and stabilization of dynamic properties are listed here, as well as some indicators in the sense of the possibility of achieving various effects. The most developed type of adaptive control systems are adaptive optimal control systems, which combine high adaptability to operating conditions with optimization of motion control and energy consumption for control in the sense of a given criterion.

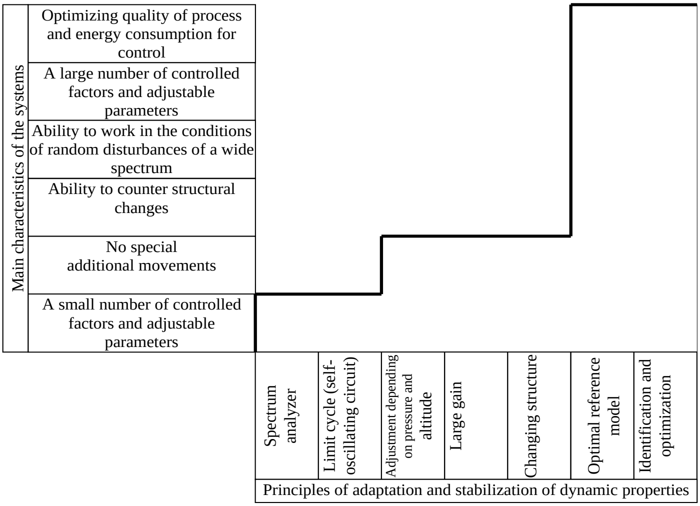


Fig. 4. Comparison of various principles of adaptation (or reduction of sensitivity to changes in aircraft parameters) of SAC

One of such systems is an adaptive control system built using Lyapunov functions and a reference model. This system is designed so that the output signal of the object eventually corresponds to the output signal of a predetermined model that has the desired characteristics [11]. Such a system must be asymptotically stable, i.e. the controlled system eventually follows the reference model with zero error. Moreover, transitional processes at the stage of adaptive control or learning have guaranteed limits. The block diagram of adaptive control with the reference model is shown in Figure 5. Such systems are closer to dual control systems in terms of goals and control results, winning over them in matters of realizability and losing to them in the organization of the interaction of state parameters assessment processes with the control process. It is expected that the implementation of the adaptive optimal flight control system will make it possible to successfully solve many problems of development and mastering new generations of UAVs, as these systems allow:

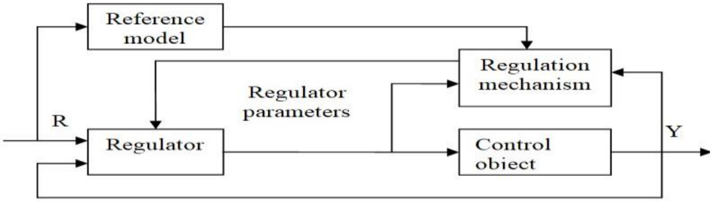


Fig. 5. Block diagram of adaptive control with reference model

– significantly expand the range of UAV application conditions;

– to ensure comprehensive optimization of the execution of functions assigned to the system;

– to increase the safety of UAV flight, including in boundary modes (modes near the borders of operational areas);

– to significantly reduce the time and material costs for the development and mastering of separate models of equipment, both due to the high, previously unavailable level of unification of the elemental base of control systems, and due to the reduction of requirements for field tests of UAVs at the initial stages of their development. Achieving these results requires solving a number of theoretical and applied problems.

2.6. Justification of choice of the optimality criterion of automatic control system of unmanned aerial vehicle

The work of DSS takes place as a rule in conditions of a priori uncertainty based on the incompleteness and inaccuracy of the input data. Also, decision-making in DSS is complicated by the stochastic nature of external influences, the lack of a correct mathematical model of functioning, the human factor, and the lack of a clear ultimate goal. In such circumstances, the risk of making ineffective decisions increases. The properties of adaptive optimal control are largely determined by the chosen optimization criterion. It should be a technical or tactical-technical criterion, expressed by the function or functional component of the object's state and management. The selection of optimality criteria is a task, which solved on the basis of an in-depth study of the controlled process. The difficulty of establishing a criterion is due to the fact that the requirements for the system are contradictory, so there is a problem of formulating a single criterion that would provide a compromise solution to the problem [12]. Thus, an adaptive optimal control system of a UAV that performs the functions of an aerial platform for telecommunication systems must simultaneously or in a controlled sequence solve majority or even all emerging flight control problems. The most universal way of implementing the optimization of control signals in the control system is the method based on the so-called combined control synthesis. The implementation of the combined synthesis of control laws leads to the fact that practically the only control input of the adaptive optimal control system is the task of a specific form of optimality criterion. These criteria determine the purpose and quality of the controlled movement. Optimality criteria, formulated in advance, determine the measure, based on which the control algorithm chooses the most favorable way for the object to achieve the given state $x_{зад}$. The structure of the control operator depends on the method of setting the control goal, minimized criteria, and the choice of the optimization method. From an engineering point of view, it seems natural to construct optimality criterion that directly take into account direct indicators of the quality of the control process [13]. These indicators (fixed error, adjustment time, re-adjustment, oscillation magnitude, period of oscillation, etc.) are physically the most understandable and have clear boundaries of permitted values. However, indirect quality indicators, which, as a rule, are easier to calculate and more convenient in analytical studies, have become more widespread in the methods of designing of control systems. These include root, frequency and integral indicators. Indirect indicators, as is known [14], are related to direct indicators, but the nature of the dependencies in most cases has not yet been revealed, which reduces the "transparency" of the requirements formulated with their help. Each of the named groups of indirect quality indicators is adequate to different methods of synthesis of optimal control systems that provide the best values of the selected indicators. For the synthesis method of controlling objects whose dynamic properties are specified in the state-space representation, integral quality indicators are the most convenient [15]. The specific form of the selected integral indicator of the quality of controlled movement is closely related to the synthesis method.

Since UAV which performs the functions of an aerial platform must be in the air for a significant period of time, its control system must minimize the energy consumption associated with maneuvering and controls. Therefore, in the general case, it is advisable to choose the form of the selected integral quality indicator as the functionals forms defined on all possible trajectories x(t) in X^n for $\forall t \in [t_0, t_K]$.

$$I = V_{ЗАД}[x(t_K), t_K] + \int_{t_0}^{t_K} Q(x,t)dt + \int_{t_0}^{t_K} L(u,t)\,dt \qquad (1)$$

Here $V_{ЗАД}$, Q and L are given functions of the specified arguments that satisfy some conditions. The first addendum in (1.1) is called the terminal term of the functional. It determines the contribution to the functionality of the final (at the moment of time t_K) state of the UAV. The second addendum in (1.1) represents an integral assessment of the quality of the transient process of the control object on the interval $[t_0, t_K]$. The third addendum is an integral estimate of the "costs" of the control signals on the interval $[t_0, t_K]$. If the functions $V_{ЗАД}$, Q and L of the functional (1) are positive definite in X^n and U^m, and their only zero values correspond to the state of the UAV, which is required by the conditions of the problem, then the optimality of the synthesized control can be understood in the sense of achieving the minimum of the functional (1). The smaller the I, the better the system and the smaller the control influences needed to achieve it. That means, by minimizing the quality criterion (1.1), we thereby minimize the energy costs required to control the UAV. Nodes of the mobile component of the military communications network must quickly adapt to frequent changes in topology and traffic levels, and also efficiently use limited network resources. In such conditions, it is impossible to ensure information exchange with the given quality without an effective control system of the mobile component, part of which is the control subsystem of the network of unmanned aerial vehicles. In work [16], the classification of control tasks of network of unmanned aerial vehicles is considered in detail, one of them is the task of deploying UAVs and controlling their flight. The deployment stage consists in launching a given number of UAVs and controlling their flight to the given barrage areas. At the same time, the tasks of the deployment stage (replanning of the topology) of the network of UAVs can be performed at the stage of operational control in case of significant changes to the CN (its damage, the introduction of new troop groups, etc.). Flight control of the UAV and the operation of its on-board systems is carried out from the network control center. At the stage of operational control, the state of CN networks is constantly assessed according to the accepted performance criteria, and actions are taken (according to the plan and the actual situation) to maintain performance indicators within the given limits or to optimize them. Tasks of operational management (as opposed to planning tasks) are solved in a mixed way (centralized/decentralized) in real time, and in terms of content they are repeated many times. According to [16], the primary goal of the network of UAVs topology control is to ensure the connectivity of all (certain zones) wireless episodic networks or priority subscribers. Therefore, as the main criterion for the effectiveness of UAV network topology control, the authors propose to choose wireless episodic network connectivity, and to transfer others to the category of restrictions. It should be noted that the definition of the optimal from the point of view of fuel consumption of the UAV trajectory change to given points in space and optimal from the point of view of energy consumption of control laws for stabilizing the aerial platform on the given trajectory while maintaining the restrictions that are imposed to ensure the quality criteria of the functioning of telecommunication systems have not yet been considered. Although the flight control subsystem is one of the main ones in the creation of telecommunication systems based on aerial platforms. Therefore, an urgent scientific task arises - scientific and methodological justification and development of a method for solving the problem of operational synthesis of the optimal energy consumption control law of a light-class UAV, which performs the functions of the aerial platform for telecommunication systems, depending on the change in the external environment (meteorological conditions, relief of the area, needs of spatial re-configuration of network) and support with high precision parameters of wireless episodic networks. To solve specified general scientific problem, it is necessary to set and solve a set of partial scientific and methodological research problems:

– analysis of the operation of wireless episodic networks based on UAV and derivation of the main limitations on the spatial movement of the aerial platform

– formulation of optimality criteria;

– development of a mathematical model of the object;

– determination of the optimal maneuvering trajectory of the UAV from the point of view of fuel consumption

– synthesis of optimal control laws;

– development of adaptation algorithms (adjustment) of control laws according to the modes of operation of the facility;

– reproduction and analysis of the received laws with the help of control digital computing machines (DCM).

The combination of the last four stages allows you to create DCM based a control system, which implement the synthesis of optimal control and control itself almost simultaneously during the operation of the object.

3. Conclusions

Thereby, the development of a method for solving the problem of the operational synthesis of the optimal energy consumption control law of a light class UAV, which performs the functions of an aerial platform for telecommunication systems on the basis of optimal adaptive control at the current stage, is an urgent scientific task, which is important for national economic and defense. Research in future works will be devoted to the solution of this problem and the composing sub-problems.

4. References

[1] P. Kirchu, Method of adaptive synthesis of the optimal energy and energy consumption law of the Internet with a telecommunication aerial platform (Systems and processes), Ph.D. thesis, National Aviation University, (2012).

[2] D. Lundstrom, Aircraft Design Automation and Subscale Testing - With Special Reference to Micro Air Vehicles. Linköping Studies in Science and Technology, Dissertations No. 1480. (2012).

[3] S. Gudmundsson, General Aviation Aircraft Design: Applied Methods and Procedures. Butterworth-Heinemann, 2021.

[4] D. Lundstrom, K. Amadori, P. Krus, Distributed Framework for Micro Aerial Vehicle Design Automation, in: 46th AIAA Aerospace Sciences Meeting and Exhibit, Reno, Nevada, 2008, pp. 1-10. Reston, Virginia. DOI:10.2514/6.2008-140.

[5] O. Dantsker, S. Imtiaz, M. Caccamo, Electric Propulsion System Optimization for Long-Endurance and Solar-Powered Unmanned Aircraft, in: AIAA Propulsion and Energy 2019 Forum, Indianapolis, 2019, pp. 1-11. DOI: 10.2514/6.2019-4486.
[6] O. Dantsker, M. Theile, M. Caccamo, A High-Fidelity, Low-Order Propulsion Power Model for Fixed-Wing Electric Unmanned Aircraft, in: 2018 AIAA/IEEE Electric Aircraft Technologies Symposium, Cincinnati, Ohio, 2018, pp. 1-14. DOI:10.2514/6.2018-5009.
[7] P. Sforza, Commercial Airplane Design Principles. Elsevier Science & Technology Books. Elsevier, Oxford, 2014.
[8] K. Amadori, D. Lundström, P. Krus, Automated design and fabrication of micro-air vehicles. J. Aerosp. Eng 226(10) (2011) 1271–1282. DOI:10.1177/0954410011419612,
[9] P. Sforza, Manned Spacecraft Design Principles. Elsevier Science & Technology Books. Elsevier, Oxford, 2015.
[10] M. Aliofkhazraei, A.S.H. Makhlouf, Handbook of Materials Failure Analysis with Case Studies from the Aerospace and Automotive Industries. Elsevier Science & Technology Books. Elsevier, Oxford, 2015.
[11] D. Lundström, K. Amadori, P. Krus, Automation of Design and Prototyping of Micro Aerial Vehicle, in: 47th AIAA Aerospace Sciences Meeting including The New Horizons Forum and Aerospace Exposition, 2009, pp. 1-16.DOI:10.2514/6.2009-629
[12] I. Moir, A. Seabridge, Design and Development of Aircraft Systems. John Wiley & Sons, Ltd, Chichester, UK, 2012. DOI:10.1002/9781118469156.
[13] O. Lysenko, O. Tachynina, O. Guida, O. Furtat, S. Furtat, I. Sushyn, Engineering Methodology of Express Synthesis of Digital Automatic Control System for Precision Mechatronic Devices of Mobile Sensor Networks, in: Proceedings of Vernadsky National University of Crimea. Technical Sciences Series. Vol. 33(72), 5, 2022, pp. 11-19.
[14] O. Lysenko, O. Tachynina, S. Ponomarenko, O. Guida, Optimal Branched Trajectories Theory. Igor Sikorsky KPI, Kyiv, 2023.
[15] M. Hassanalian, A. Abdelkefi, Classifications, applications, and design challenges of drones: A review. Progress in Aerospace Sciences. Vol. 91., 2017 pp. 99-131.

УДК 519.85

ПІДХІД ДО ВИЗНАЧЕННЯ ВПЛИВУ ПАРАМЕТРІВ ПРОСТОРОВОГО РУХУ БПЛА НА ХАРАКТЕРИСТИКИ

Лисенко О.[а], Dc.Sci. Тачініна О.[б], Dc.Sci. Кирчу П.[c], Ph.D. Пономаренко С.[а], Ph.D. Кутєпов В.[б], Сушин І.[а]

[a] Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Україна
[b] Національний авіаційний університет, Україна
[c] Компанія «ЮТАС», Україна
E-mail: tachinina5@gmail.com, lysenko.a.i.1952@gmail.com, guydasg@ukr.net, fkirchu@gmail.com, novikov1967@ukr.net, rubin268@ukr.net

Анотація. *Ключовим елементом новітніх систем широкосмугового бездротового доступу, що працюють в умовах мегаполісу, гірської або пересіченої місцевості, а також в умовах бойових дій і надзвичайних ситуацій, є висотна підйомна платформа зі спеціальним телекомунікаційним обладнанням, розташована на це. Використання для таких систем малогабаритних безпілотних літальних апаратів (БПЛА) авіаційного типу дозволяє досягти наступних переваг: відносно низька вартість впровадження, малі габарити, короткий період часу для розгортання систем зв'язку в заданій місцевості. У розділі розглядаються галузі реалізації та принципи структурно-функціональної побудови телекомунікаційних систем на базі висотних літальних платформ. Проведені дослідження в галузі розробки систем керування БПЛА дозволили визначити основні переваги та недоліки новітніх підходів до синтезу законів керування польотом та умови їх застосування. Вибір оптимальної адаптивної системи керування автовишкою дав змогу обрати критерієм оптимізації квадратичний функціонал. Це дасть змогу забезпечити мінімізацію енергоспоживання при збереженні необхідних якісних показників функціонування телекомунікаційної системи, в умовах зміни параметрів об'єкта керування.*

Ключові слова: *БПЛА, літальна платформа, телекомунікаційна мережа, бездротова епізодична мережа*

UDC 681.51:004.896

FUZZY AUTOMATIC CONTROL SYSTEM OF MOBILE ROBOTIC PLATFORM'S ADHESION

Zheng Y.[1], Dr.Sci. Kozlov O.[2], Dong C.[1], Ph.D. Kondratenko G.[2], Skakodub O.[2]

ORCID: 0000-0003-4690-098X, 0000-0003-2069-5578, 0000-0003-3529-6529, 0000-0002-8446-5096, 0000-0002-5643-775X

[1] Yancheng Polytechnic College, China,

E-mail: and_bsb@126.com, dongchengjian1988@126.com

[2] Petro Mohyla Black Sea National University, Ukraine,

E-mail: kozlov_ov@ukr.net, halyna.kondratenko@chmnu.edu.ua, aleksandrskakodub1996@gmail.com

Abstract. *The chapter is focused on modern aspects of mobile robotics, namely to the universal mobile robotic platforms that can be used in various technological environments and conditions to perform different technical tasks at industrial facilities and enterprises. These mobile platforms are capable of moving on horizontal surfaces with complex terrain as well as climbing vertically on sheer walls and ceilings and are the universal autonomous means of carrying out complex operations in hard-to-reach and dangerous places for humans. One of the biggest challenges in operating these robotic platforms is proper adhesion control when moving on inclined and vertical planes. In this study, the authors performed the design and study of the fuzzy automatic control system of adhesion that provides efficient and reliable fastening and movement of the platform on inclined surfaces of various types. The proposed system is based on fuzzy inference engine of Mamdani type that allows determining the necessary adhesion force of the robotic platform implementing expert knowledge for its safe and efficient use at various angles of the surface inclination. The system's performance is verified by computer simulation.*

Keywords: *mobile robotics, universal mobile robotic platform, intelligent adhesion control, fuzzy automatic control system, computer simulation*

1. Introduction

In recent years, the introduction of robotic systems and complexes in various spheres of human activity has become increasingly widespread [1-3]. In modern life, different types of robots are involved in a huge number of processes, ranging from closed production cycles of heavy industry and metallurgy to customer service and cargo delivery. The use of various robotic solutions provides enormous benefits, in particular, eliminating expensive human labor, significantly increasing productivity, accuracy and speed of technological operations, reducing risks to the life and health of people when performing tasks in hazardous and harmful environments, eliminate errors caused by the human factor and operator fatigue, etc. [4, 5].

A separate fairly widely used class of robotic systems are mobile robots (MR) and robotic complexes [6, 7]. They allow performing the tasks of monitoring, inspection, reconnaissance, the movement of equipment and tools to perform complex labor-intensive technological operations in hard-to-reach places, and are especially effective at functioning in the fully automatic mode [8, 9]. For example, at the moment, the mobile welding robot has been developed and successfully applied, that can achieve automatic large fillet welding seam tracking in narrow spaces [10]. In turn, MRs for painting [11] and removing paint [12] from various surfaces are quite effective. Also, widespread are robots for inspection [13] and complex cleaning [14] of the bodies of various objects, and others. Universal mobile robotic platforms (UMRP) have even greater efficiency [15]. Due to the modular structure, these robotic objects can simultaneously have different types of propulsion devices and different sets of technological tools. This allows them to be used to perform a wide range of diverse tasks in various environments and conditions, for example, on a horizontal surface with complex terrain, above and below water, on vertical and inclined surfaces of various types, etc. [16]. As a result, due to such versatility, the speed of performance is significantly increased and the cost of performing various technological operations is reduced. However, at the same time, the creation of such types of UMRPs also poses rather difficult tasks for developers to design highly efficient intelligent control systems. These control systems should also have a modular structure to automate the control processes of various types of propulsion devices (for moving on a horizontal and inclined surface, underwater and above water) and various working tools (for monitoring, inspection, painting, cleaning, welding, etc.). It is most expedient to develop such control systems based on the principles of artificial intelligence [17], which is confirmed by a number of published studies [18, 19]. So, fuzzy control is used quite effectively for mobile robot navigation in an unknown environment when avoiding random obstacles [20] and for a caterpillar robot capable of moving along inclined ferromagnetic surfaces [21]. Neural network control is also widely used, for example, for inspection mobile robots [22] and others. Of particular note is the adhesion force control system (AFCS) when the platform moves on various types of inclined surfaces. This system is the linchpin for ensuring safe and efficient movement in such challenging terrains, which is confirmed by a number of recent studies [23, 24]. Without precise control over adhesion, the platform could face significant risks including slippage, loss of traction, or even potential accidents. Moreover, the system directly impacts the overall efficiency of operations, as it regulates the force necessary for the platform to securely adhere to the surface. Proper adhesion control not only ensures stability but also enables the robotic platform to perform a wide array of tasks in otherwise hard-to-reach and hazardous environments. Thus, the adhesion force control system serves as a critical component for unlocking the full potential of a universal robotic platform, making it an indispensable feature for a diverse range of applications. Moreover, the correct control of the adhesion force can significantly increase the efficiency of the entire motion control system due to the rational distribution of loads. In particular, the adhesion force control systems for various robots have been developed based on advanced control principles, namely predictive [25] and fuzzy [26], neuro-fuzzy [9]. The systems in [25, 26] make it possible to implement smooth adhesion control in various conditions (when changing the driving torque and surface friction), however, do not take into account the simultaneous uncertain change in the inclination angle of the working surface. In turn, the system in paper [9], on the contrary, takes into account the uncertain change in the angle of inclination and the friction coefficient of the surface, but does not consider the different values of the driving torques of the propulsors. At the same time, to achieve high adhesion efficiency, it is necessary to implement flexible determining and automatic control of the adhesion force depending on the coefficient of friction, driving torque and the angle of inclination of the working surface. Thus, the main aim of this chapter is the development and research of the fuzzy system for adhesion automatic control of the universal mobile robotic platform with flexible determining the adhesion force taking into account the main parameters: coefficient of friction, driving torque, inclination angle of the working surface. The rest of this chapter is organized as follows. Section 2 presents the developed generalized structure of the hierarchical control system for the universal mobile robotic platform. In section 3, the authors discuss in detail the functional structure of the two-level adhesion force control system for the universal mobile robotic platform. In turn, the development of the tactical-level fuzzy subsystem for calculating the adhesion force set value is considered in section 4. Section 5 outlines the prospects for further optimization of the designed fuzzy subsystem for calculating the adhesion force. The chapter ends with a conclusions in section 6 and acknowledgements in section 7.

2. Generalized Structure of the Hierarchical Control System for the Universal Mobile Robotic Platform

Since the universal mobile robotic platform is a complex multi-component control plant, its control system must have a multi-level hierarchical structure. In turn, the principle of hierarchical multi-level control implies the presence of the highest, strategic, tactical and executive levels of control. The highest level represents the human operator and the human-machine interface (HMI) [27]. At this level of the control system, the human operator must make a decision on the implementation of a particular technological operation, task, movement, or maneuver of the UMRP based on the analysis of the environment and the situation, assessment of current operating conditions, external disturbances, etc. [28]. Moreover, to make such decisions the operator may implement preliminary simulations of certain situations based on the embedded simulation models to predict the UMRP's behavior and the state of the environment. At the strategic level of control, after receiving certain commands (control goals) from the highest level it is necessary to plan technological operations, tasks, or movements and their transformation into certain sequences of elementary actions (subtasks) [28, 29, 30]. At this level, various control algorithms can be used depending on the specific technological operation being performed. Therefore, the strategic level forms the commands for the tactical level to perform specific actions or basic operations. In turn, the main task of the tactical control level is to transform the control commands of the strategic level into the control programs that define the laws of coordinated functioning or movement of executive mechanisms and actuators of the executive level of control. The given programs define sequences of the set values of the generalized controlled coordinates of the UMRP's main executive mechanisms [28]. As for the executive level of control, it consists directly of executive mechanisms and actuators, sensory system (SS), as well as of automatic control subsystems, which due to the corresponding control impacts work out the set values of the platform's generalized controlled coordinates, that come from the tactical level [28]. In turn, the SS is used to receive feedback and to obtain all available information about the state of the UMRP and the environment. The generalized functional structure of the hierarchical control system for the UMRP is presented in Fig. 1, where the following notations are adopted: CATO1, CATO2, …, CATOl are the control algorithms of the 1st, 2nd, …, lth technological operations; CMPD1, CMPD 2, …, CMPDn are the control modules of the 1st, 2nd, …, nth propulsion devices; CMTT1, CMTT 2, …, CMTTm are the control modules of the 1st, 2nd, …, mth technological tools; USL is the vector of control signals for the strategic level; UTL is the vector of control signals for the tactical level; UEL is the vector of control signals for the executive level; USS is the vector of sensory system outputs; XR is the vector of controlled and technological coordinates of the UMRP.

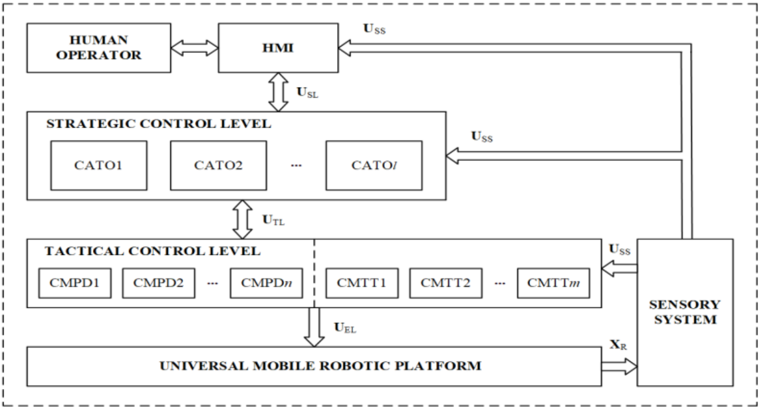


Fig. 1. Generalized functional structure of the hierarchical control system for the universal mobile robotic platform

In this control system the operator uses a specialized HMI to transmit the control signals USL to the strategic control level and receive signals USS about the state of the platform and the environment from the sensory system. The strategic level has a set of l control algorithms (CATO1, CATO2, …, CATOl) for implementing the control of a number of technological operations (inspection, welding, painting, ultrasonic diagnostics, rust removal, cleaning, etc.) that can be performed using this robotic platform. In turn, various combinations of propulsion devices and technological tools of the mobile platform can be used to perform these technological operations. Moreover, to perform new types of technological operations, appropriate new control algorithms can be added in this system to the strategic level of control. For direct control of the different propulsion devices (wheeled, caterpillar, walking, propeller, gravity type, etc.) and technological tools (video cameras, welding machines, cleaning cutters, manipulators and others) in the process of performing various operations, there are corresponding control modules (CMPD1, CMPD2, …, CMPDn, CMTT1, CMTT2, …, CMTTm) at the tactical control level in this system. When adding new types of propulsion devices or technological tools to the UMRP, the appropriate control modules must be previously designed and added to the tactical control level of the system. In turn, these control modules are control systems with a complex structure and carry out the coordinated control of all executive mechanisms, drives and actuators that are parts of certain propulsion devices and technological tools. Finally, to control the variables of individual drives and actuators at the executive level (located directly on the UMRP in Fig. 1) of this system, there are separate stabilization and automatic control subsystems, which are slave control systems for the tactical-level control modules. Among the many working tools and propulsion devices of the mobile robotic platform, the adhesion device (AD) deserves special attention. The use of this device gives the opportunity to significantly expand the range of tasks and technological operations, as it becomes possible to perform various work in hard-to-reach places on inclined and vertical surfaces (sheer walls and ceilings). In turn, the adhesion device, depending on the tasks and operating conditions, can be implemented on the basis of various physical principles, for example, propeller type, vacuum, magnetic, electromagnetic, and others. Thus, for the effective application of the adhesion devices of various types on the UMRP in various operating modes, it is necessary to develop an appropriate universal control module. Next, we consider the development of a functional structure, control algorithms and the main components of the AD control module in the form of a specialized adhesion force control system.

3. Functional Structure of the Two-level Adhesion Force Control System for the Universal Mobile Robotic Platform

The main task of the adhesion device of the mobile robotic platform is to create the necessary adhesion force to an inclined or vertical surface for the safe and efficient movement of this platform while simultaneously performing the necessary technological operation. To safely hold the UMRP on an inclined or vertical surface, the maximum possible adhesion force could be provided in all operation modes, however, in this case, there will be a maximum energy consumption of the adhesion device and a sufficiently large additional resistance to the propulsion devices. Together, this will significantly reduce the overall efficiency of the UMRP and the performance of a particular operation. Thus, for the most efficient use of the adhesion device with energy saving, it is necessary to provide flexible control of the adhesion force depending on the various modes of movement of the platform and the performance of various technological operations, as well as on many other factors. At the same time, the AFCS must determine the required (set) value of the adhesion force and ensure its automatic maintenance (stabilization). In turn, the set value of the adhesion force is significantly affected by such parameters as the angle of inclination and the friction coefficient of the working surface, the total mass of the robotic platform together with the technological tools, as well as the traction force of the propulsion devices. Moreover, the structure of this control system must be universal for any type of the adhesion device. Taking into account the above conditions and particular features, as well as the complexity of the mathematical description of the processes of the platform moving along inclined surfaces of various types, it is advisable to develop the adhesion force control system based on the principles of artificial intelligence [31, 32]. In addition, this system should consist of two levels of control: tactical and executive. At the tactical level, the set value of the adhesion force of the platform to the surface should be determined based on the values of the angle of inclination of the working surface, the coefficient of friction and the traction force of the propulsion devices. In turn, at the executive level, the automatic control of the adhesion force should be performed, that is, the stabilization of the set value under the influence of various disturbances. Analyzing various methods and approaches of artificial intelligence, it can be concluded that it is most expedient to develop the tactical-level subsystem for determining the set value of the adhesion force on the basis of fuzzy logic [33, 34]. Fuzzy systems make it possible to effectively generalize expert information and experimental data, formalize the mechanisms of human thinking, form linguistic models of complex processes, and approximate nonlinear multidimensional dependencies [35, 36]. In turn, the executive-level subsystem for the adhesion force automatic control can be designed based on different types of intelligent controllers, in particular, neural network [37, 38], fuzzy [39] or neuro-fuzzy [9].Taking into account all of the above, the functional structure of the two-level adhesion force control system for the UMRP is formed, which is shown in Fig. 2.

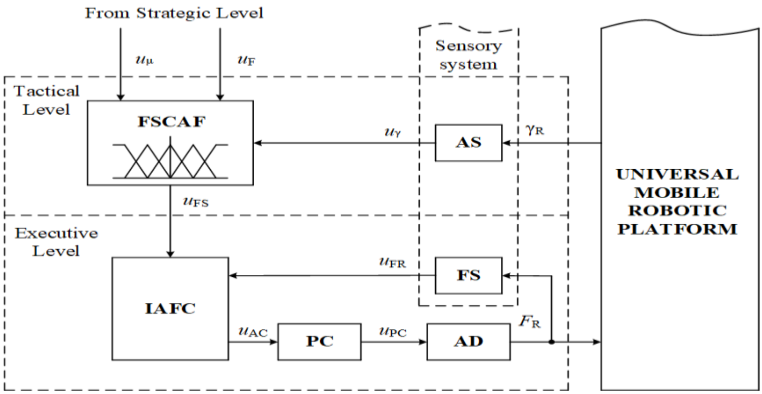


Fig. 2. Functional structure of the two-level adhesion force control system for the universal mobile robotic platform

In Fig. 2, the following designations are adopted: FSCAF is the fuzzy subsystem for calculating the adhesion force set value; IAFC is the intelligent adhesion force controller; AS is the angle sensor; FS is the force sensor; PC is the power converter; uμ is the signal corresponding to the current value of the coefficient of friction of the contact parts of the platform and the working surface; uF is the signal corresponding to the current value of the traction force of the UMRP propulsion devices; uγ is the signal corresponding to the current value of the angle of inclination of the working surface; uFS is the FSCAF output signal which corresponds to the set value of the adhesion force; uFR is the FS output signal which corresponds to the real value of the adhesion force; uAC is the IAFC output control signal; uPC is the PC output signal; γR is the current value of the angle of inclination of the working surface; FR is the real value of adhesion force. As can be seen from Fig. 2, at the tactical level, the FSCAF determines the required (set) value of the adhesion force based on the signals corresponding to the current value of the coefficient of friction uμ, the current value of the traction force of the UMRP propulsion devices uF and the current value of the angle of inclination of the working surface uγ. In turn, the signals uμ and uF are given from the strategic level of control, and the signal uγ comes from the sensor for measuring the angle of inclination of the working surface. The current value of the coefficient of friction of the contact parts of the platform and the working surface is determined experimentally or pre-set by the operator depending on the type and material of the working surface. The current value of the traction force of the UMRP propulsion devices comes from the control system of the platform propulsion devices. As a result, the signal which corresponds to the set value of adhesion force uFS is determined by the FSCAF based on fuzzy inference engine and a pre-designed rule base. At the executive level, the stabilization of the set value of the adhesion force is performed using the intelligent adhesion force controller. In turn, the IAFC the compares the signal uFS from the FSCAF with the force sensor signal uFR and, using the embedded intelligent algorithm, performs automatic control of the UMRP's adhesion force. The IAFC output control signal uAC is amplified by a power converter and directly fed to the adhesion device. Moreover, the intelligent controller must be previously adjusted or trained for a specific adhesion device based on the simulation or experimental models. Next, we consider the development of the tactical-level fuzzy subsystem for determining the necessary adhesion force of the UMRP in more detail.

4. Development of the Tactical-level Fuzzy Subsystem for Calculating the Adhesion Force Set Value

The fuzzy subsystem for calculating the set value of the adhesion force (Fig. 3) has three inputs (uγ, uμ, uF) and one output (uFS). It is advisable to develop this fuzzy system on the basis of a Mamdani-type fuzzy inference mechanism. In turn, Mamdani's fuzzy inference mechanism includes sequential execution of the following stages: fuzzification, aggregation, activation, accumulation and defuzzification [40, 41, 42]. At the stage of fuzzification, the instantaneous numerical values of the input variables are mapped to the corresponding fuzzy term sets with the calculation of the membership degree values. In this case, for the variable uγ, corresponding to the angle of inclination of the working surface, it is advisable to choose 6 linguistic terms with triangular-type membership functions: Z – zero; S – small; LM – less than middle; M – middle; L – large; VL – very large.

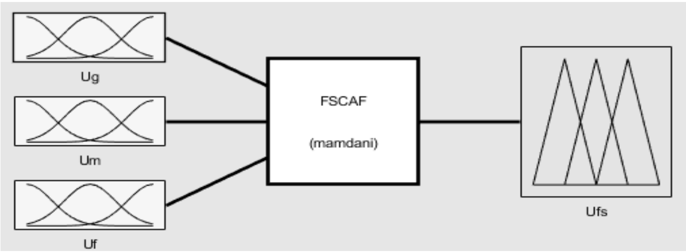


Fig. 3. Structure of the fuzzy subsystem for calculating the set value of the adhesion force

Also, this variable can vary in the range from 0 to 180 degrees. Moreover, the variable uγ is defined by the value of angle γ'R that is determined by the current value of the angle of the working surface inclination γR using the following dependency

$$\gamma'_{\mathrm{R}} = \begin{cases} \gamma_{\mathrm{R}}, \text{ at } \gamma_{\mathrm{R}} \le 180°; \\ 360° - \gamma_{\mathrm{R}}, \text{ at } \gamma_{\mathrm{R}} > 180°. \end{cases} \quad (1)$$

The appearance of the linguistic terms of the variable uγ with the set parameters is shown in Fig. 4.

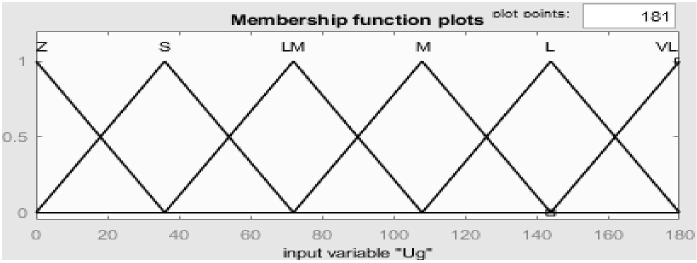


Fig. 4. Linguistic terms of the variable uγ with the set parameters

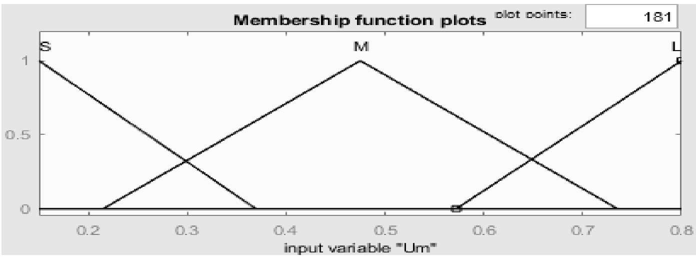
The membership function of linguistic terms of triangular type on the example of this variable is represented by the expression (2)

$$\mu(u_\gamma) = \begin{cases} 0, \text{ at } u_\gamma \le a \text{ or } u_\gamma \ge c; \\ \dfrac{u_\gamma - a}{b-a}, \text{ at } a < u_\gamma \le b; \\ \dfrac{c-u_\gamma}{c-b}, \text{ at } b < u_\gamma < c; \end{cases}$$

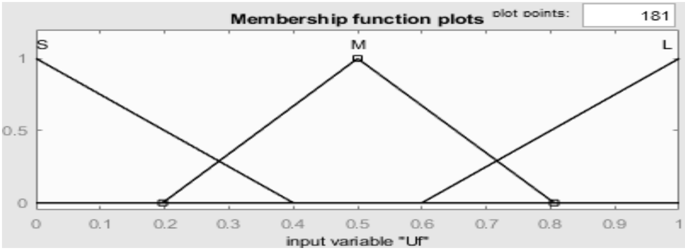
$$a \le b \le c, \tag{2}$$

where a, b and c are the customizable parameters of the membership function.

In turn, for the variables uμ and uF, that correspond to the current values of the friction coefficient and traction force of the propulsion devices, 3 linguistic terms with triangular-type membership functions are chosen: S – small; M – middle; L – large. The appearance of the linguistic terms of these variable with the set parameters is shown in Fig. 5, a, b.



a



b

Fig. 5. Linguistic terms with the set parameters of the variables: a) uμ; b) uF

The variable uμ can vary in the range from 0.15 to 0.8, and the variable uF is set in relative units (from 0 to 1) from the maximum value of the traction force. As for the system's output variable uFS, it is advisable to use 9 linguistic terms with triangular-type membership functions for it. They are: Z – zero; S – small; LM – less than middle; M – middle; MM – more than middle; L – large; VL – very large; E – extreme. In turn, this variable is set in relative units (from 0 to 1) from the maximum value of the adhesion force developed by the adhesion device. As a rule, the maximum adhesion force value is defined as 7...10FP, where FP is the total UMRP weight with equipment. The appearance of the linguistic terms of this output variable with the set parameters is shown in Fig. 6.

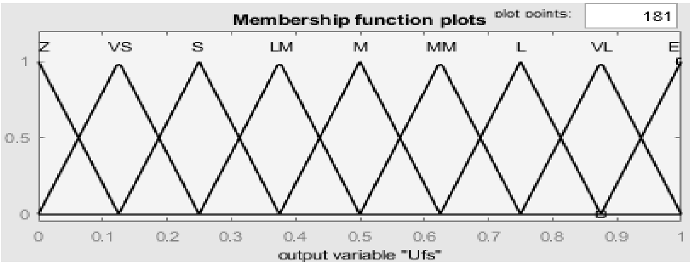


Fig. 6. Linguistic terms of the output variable uFS with the set parameters

The production rules of the rule base (RB) for the given Mamdani-type system are given as follows:

$$\text{IF “}u_\gamma = A_1\text{” AND “}u_\mu = B_1\text{” AND “}u_\text{F} = C_1\text{” THEN “}u_\text{FS} = D_1\text{”}, \quad (3)$$

where A1, B1, C1 and D1 are certain linguistic terms of the given variables. The developed rule base of the fuzzy subsystem for determining the necessary adhesion force of the UMRP is presented in Table 1.

The Mamdani-type fuzzy inference engine consists of sequential execution of the stages of aggregation, activation and accumulation [40, 41]. At the aggregation stage, the degree of truth of the conditions for each of the rules of the fuzzy inference system is determined [43]. For this, the values of the membership functions of the linguistic terms of the variables obtained at the stage of fuzzification, which make up the antecedents of fuzzy production rules, are used. In turn, finding the degrees of membership of antecedents is done using the t-norm based on the “min” operation. The activation stage is the process of finding the degree of truth of each of the elementary logical subconclusions (expressions) that make up the consequents of all fuzzy production rules [44]. In turn, for the output signal of the system, at this stage, truncated membership functions for subconclusions of the rules are found based on the “min” operation. At the accumulation stage, the truncated membership functions found at the previous stage are combined to obtain the final fuzzy subset of the output variable based on the “max” operation.

Table 1

Rule base of the fuzzy subsystem for calculating the adhesion force set value

Rule number	Input variables			Output variable
	u_F	u_μ	u_γ	u_{FS}
1	S	S	Z	VS
2	S	S	S	LM
3	S	S	LM	MM
4	S	S	M	L
5	S	S	L	M
6	S	S	VL	S
7	S	M	Z	Z
8	S	M	S	S
9	S	M	LM	M
10	S	M	M	MM
11	S	M	L	LM
12	S	M	VL	S
13	S	L	Z	Z
14	S	L	S	VS
15	S	L	LM	S
16	S	L	M	LM
17	S	L	L	S
18	S	L	VL	S

Table 1 (continue)

19	M	S	Z	S
20	M	S	S	M
21	M	S	LM	L
22	M	S	M	VL
23	M	S	L	MM
24	M	S	VL	LM
25	M	M	Z	VS
26	M	M	S	LM
27	M	M	LM	MM
28	M	M	M	L
29	M	M	L	M
30	M	M	VL	LM
31	M	L	Z	VS
32	M	L	S	S
33	M	L	LM	LM
34	M	L	M	M
35	M	L	L	LM
36	M	L	VL	LM
37	L	S	Z	LM
38	L	S	S	MM
39	L	S	LM	VL
40	L	S	M	E
41	L	S	L	L
42	L	S	VL	M
43	L	M	Z	S
44	L	M	S	M
45	L	M	LM	L
46	L	M	M	VL
47	L	M	L	MM
48	L	M	VL	M
49	L	L	Z	S
50	L	L	S	LM
51	L	L	LM	M
52	L	L	M	MM
53	L	L	L	M
54	L	L	VL	M

In turn, the defuzzification stage is a process of transition from the membership function of the output linguistic variable $\mu_{\Sigma}\left(u^{*}_{\mathrm{FS}}\right)$ to its clear (numerical) value uFS [42]. For fuzzy inference of the Mamdani type, the numerical value of the output signal is calculated by the center of gravity method based on the dependence (4). The visualization of the calculation of the necessary adhesion force value using the developed fuzzy subsystem based on the available rules is shown in Fig. 7.

$$u_{\mathrm{FS}} = \frac{\int u_{\mathrm{FS}}^{*} \mu_{\Sigma}\left(u_{\mathrm{FS}}^{*}\right) du_{\mathrm{FS}}}{\int \mu_{\Sigma}\left(u_{\mathrm{FS}}^{*}\right) du_{\mathrm{FS}}}. \tag{4}$$

In this case, the calculation of the adhesion force value is performed for the following values of the input variables: u_{γ} = 90; u_{μ} = 0.326; u_{F} = 0.271. In turn, at these values of the inputs, the current value of the adhesion force signal u_{FS} is 0.661.The characteristic surfaces of the developed fuzzy subsystem for calculation of the necessary adhesion force value are presented in Fig. 8-10. In particular, the dependences $u_{\mathrm{FS}} = f(u_{\gamma}, u_{\mu})$ at u_{F} = 0.1 and at u_{F} = 0.9 are shown in Fig, 8, *a* and *b*, respectively. Moreover, the dependences $u_{\mathrm{FS}} = f(u_{\gamma}, u_{\mathrm{F}})$ at u_{μ} = 0.1 and at u_{μ} = 0.9 are given in Fig, 9, *a* and *b*, respectively. And, finally, the dependences $u_{\mathrm{FS}} = f(u_{\mu}, u_{\mathrm{F}})$ at u_{γ} = 30 and at u_{γ} = 90 are shown in Fig, 10, *a* and *b*, respectively. As can be seen from Fig. 7-10, the developed tactical-level fuzzy subsystem is quite efficient in determining the required value of the adhesion force in different operation modes based on the signals corresponding to the current value of the coefficient of friction u_{μ}, the current value of the traction force of the UMRP propulsion devices u_{F} and the current value of the angle of inclination of the working surface u_{γ}. This system can be successfully applied to determine the force for various types of the adhesion devices.

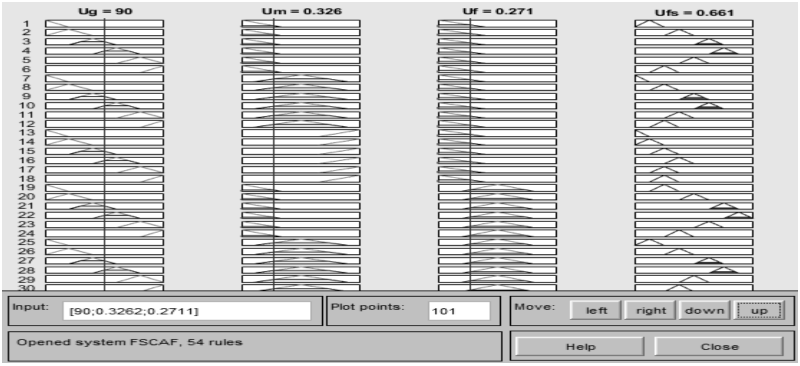
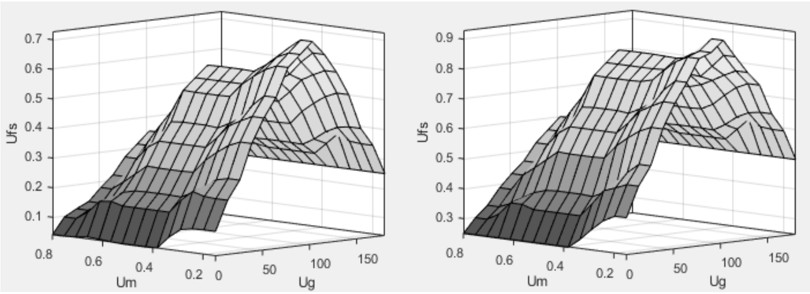
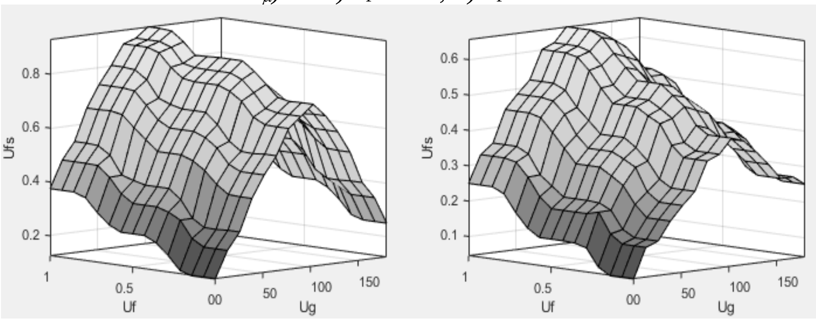


Fig. 7. Visualization of the calculation of the necessary adhesion force value using the developed fuzzy subsystem



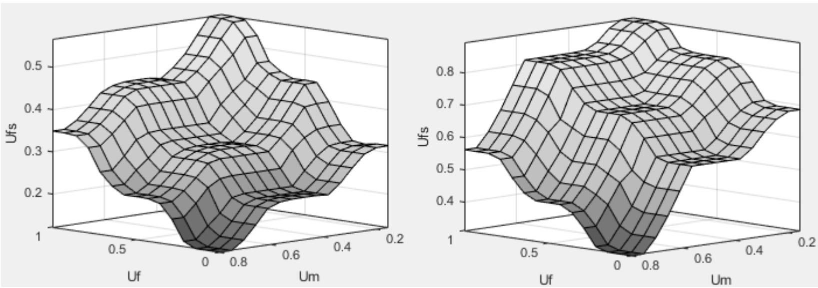
a b

Fig. 8. Characteristic surfaces of the developed fuzzy subsystem $u_{FS} = f(u_\gamma, u_\mu)$ at: a) $u_F = 0.1$; b) $u_F = 0.9$



a b

Fig. 9. Characteristic surfaces of the developed fuzzy subsystem $u_{FS} = f(u_\gamma, u_F)$ at: a) $u_\mu = 0.2$; b) $u_\mu = 0.7$



a b

Fig. 10. Characteristic surfaces of the developed fuzzy subsystem $u_{FS} = f(u_\mu, u_F)$ at: a) $u_\gamma = 30$; b) $u_\gamma = 90$

For stabilization and automatic control of the set value uFS (calculated by the FSCAF) of the UMRP adhesion force under the action of various uncertain disturbances it is planned to design the executive-level intelligent subsystem based on neural network controller [45] in further studies. If it is necessary to increase the accuracy of determining the required adhesion force of the developed system, as well as to generally improve the quality indicators of control of a universal mobile robotic platform, further optimization procedures can be applied. Next, we will briefly consider the prospects for optimization processes of the proposed fuzzy system for determining the required adhesion force.

5. Prospects for Further Optimization of Fuzzy Subsystem for Calculating the Adhesion Force Set Value

The prospects for optimizing the developed fuzzy system to determine the required adhesion force are highly promising and hold significant potential for enhancing its quality indicators. Through systematic refinement and fine-tuning, the fuzzy system can achieve higher levels of accuracy and precision in estimating the required adhesion force. This optimization process entails a thorough examination of membership functions, rule sets, and input parameters, ensuring they are finely calibrated to capture subtle variations in the operating environment. Additionally, the incorporation of advanced algorithms and machine learning techniques could further elevate the system's performance. This heightened accuracy would not only bolster the safety and stability of clamping operations but also lead to substantial improvements in overall operational efficiency and productivity. Thus, the ongoing efforts towards optimizing the fuzzy system represent a crucial avenue for advancing the state-of-the-art in adhesion force determination, with far-reaching implications for diverse industrial applications. For the proposed fuzzy system, both parametric and structural optimization can be carried out [39-41]. Herewith, the tuning of the parameters for linguistic terms membership functions and normalizing factors are considered to be the parametric optimization procedures. At the same time, the determination of proper linguistic terms number and types, mechanisms of fuzzy inference engine and defuzzification method, as well as RB antecedents and RB reduction, are structural optimization procedures. These optimization tasks can be solved sequentially with different priority or in parallel using specially selected methods and information technologies. Moreover, the analysis of recent studies indicate that progressive bioinspired methods and information technologies can be implemented quite effectively for the parametric and structural optimization of fuzzy decision support and control systems of different types [34, 42, 43]. The bioinspired intelligent methods have the following advantages and attractive features. The first one is the strong ability of local minima avoiding compared to the conventional optimization methods [41, 44, 46]. This feature is provided due to the stochastic nature of these algorithms that in practice allow prevent getting stuck in local solutions. Also, it gives the opportunity to extensively investigate the entire search space, which is multimodal, unknown, and has the complex relief when optimizing the fuzzy systems for different applications. Another important advantage is the usage of derivation-free mechanisms in the optimization process [47]. In comparison with the gradient-based approaches of optimization, the bioinspired meta-heuristic techniques do not impose any restrictions on the objective functions, start the search process from random points, and have no need to calculate the derivatives of the objective function for finding optimal solutions. Moreover, flexibility and relative simplicity can be considered as attractive features of the bioinspired methods [48, 49]. These methods can be easily adapted for solving completely different real-life optimization problems as well as for the synthesis and optimization of the structure and different parameters of high-dimensional fuzzy systems. Furthermore, these algorithms use enough simple (in terms of computational complexity) basic computational procedures in the optimization process, that simulate evolutionary concepts, social animals' behaviors, physical phenomena, etc. Finally, bioinspired intelligent methods can be effectively hybridized with local search techniques of different types which allows to rationally take advantage of global and local search strategies [50]. As a result of such combinations, specific hybrid methods are obtained that require fewer iterations to find global optima at solving optimization problems [50, 51].

6. Conclusion

In this work the development and research of the fuzzy system for adhesion automatic control of the universal mobile robotic platform is carried out. The proposed system has the two-level structure (with tactical and executive levels of automatic control) and is based on the principles of artificial intelligence that provide flexible automatic control of the adhesion force depending on the various modes of movement of the platform and the performance of various technological operations, as well as on many other factors. In particular, the tactical-level subsystem is designed on the basis of fuzzy logic and allows determining the necessary adhesion force of the UMRP to the working surface taking into account the current values of the angle of inclination of the surface, the coefficient of friction and the traction force of the propulsion devices. In turn, the designed fuzzy subsystem is implemented on the basis of Mamdani-type inference engine and has 54 production rules in the rule base compiled on the basis of expert knowledge. This gives the opportunity to ensure the most efficient use of the adhesion device with energy saving that provides reliable holding and movement of the platform on the working surface with different inclinations and characteristics in the process of performing various technological operations. The high efficiency of the developed fuzzy subsystem of the tactical level is confirmed by the simulation results obtained in the form of characteristic surfaces for various conditions. Moreover, the additional advantage of this system is that it can be successfully applied to determine the adhesion force for various types of the adhesion devices (propeller, vacuum, magnetic, electromagnetic, and others). The conducted analysis of recent studies shows that progressive bioinspired methods and approaches can be implemented quite effectively for the parametric and structural optimization of the designed fuzzy system for determining the necessary adhesion force of the UMRP. Further research should be carried out in the direction of hardware and software implementation of the UMRP’s fuzzy system and conducting full-scale experiments for comparison with the obtained modeling results.

7. Acknowledgements

This study is financially supported by the National High Level Foreign Experts Introduction Project, China (G2022014063L). We thank the Introduction of Intelligence Projects for Policy Guidance in Yancheng (YCBX2022005) and Scientific Research Project of Yancheng Polytechnic College (ygy2103).

8. References

[1] B. Siciliano, O. Khatib, Springer handbook of robotics, Springer, 2016.
[2] Y. P. Kondratenko, Robotics, automation and information systems: future perspectives and correlation with culture, sport and life science, in: A.M. Gil-Lafuente, C. Zopounidis (Eds.), Decision making and knowledge decision support system: Lecture Notes in Economics and Mathematical Systems, Switzerland, Springer International Publishing, Vol. 675, 2015, pp. 43-56.
[3] R. Duro, et al. (Eds), Advances in intelligent robotics and collaborative automation, River Publishers, Aalborg, 2015. DOI: 10.13052/rp-9788793237049
[4] E. Garcia. et al., The evolution of robotics research, in: IEEE Robotics & Automation Magazine, Vol. 14, No. 1, 2007, pp. 90-103. DOI: 10.1109/MRA.2007.339608.
[5] Y.P. Kondratenko, V.Y. Kondratenko, Advanced Trends in Design of Slip Displacement Sensors for Intelligent Robots, in: R. Duro, Y. Kondratenko (Eds.), Advances in Intelligent Robotics and Collaborative Automation, River Publishers, 2015, pp. 167-192. DOI: 10.1201/9781003337119-8
[6] Mobile Robotics: Solutions and Challenges, Proceedings of the Twelfth International Conference on Climbing and Walking Robots and the Support Technologies for Mobile Machines, O. Tosun, H. L. Akin, M. O.Tokhi, G. S. Virk (Eds.), Istanbul, Turkey, 2009.
[7] M. Derkach, D. Matiuk, I. Skarga-Bandurova, Obstacle Avoidance Algorithm for Small Autonomous Mobile Robot Equipped with Ultrasonic Sensors, in: Proceedings - 2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies, DESSERT 2020, 2020, pp. 236-241. DOI: 10.1109/DESSERT50317.2020.9125019
[8] M. Taranov, et al., Simulation of robot's wheel-mover on ferromagnetic surfaces, Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, in: Proceedings of the 2017 IEEE 9th, International Conference, Vol. 1, 2017, pp. 283-288. DOI: 10.1109/IDAACS.2017.8095091
[9] Y. P. Kondratenko, et al., Neuro-fuzzy observers of clamping force for magnetically operated movers of mobile robots, Technical Electrodynamics 5 (2017) 53-61. DOI: 10.15407/techned2017.05.053.
[10] Z. -W. Mao, J. -l. Pan, H. Zhang, Mobile welding robot system based on rotating arc sensor applied for large fillet welding seam tracking, in: 2010 Sixth International Conference on Natural Computation, Yantai, China, 2010, pp. 394-397. DOI: 10.1109/ICNC.2010.5583304.
[11] M. Sorour, RoboPainter – A detailed robot design for interior wall painting, in: 2015 IEEE International Workshop on Advanced Robotics and its Social Impacts (ARSO), Lyon, France, 2015, pp. 1-6. DOI: 10.1109/ARSO.2015.7428214.
[12] B. Ross, J. Bares, C. Fromme, A Semi-Autonomous Robot for Stripping Paint from Large Vessels, The International Journal of Robotics Research (2008) 617-626.
[13] M. Cardona, J. Cerrato, E. García, Design and Simulation of a Mobile Robot for Pipeline Inspection, in: 2022 IEEE Central America and Panama Student Conference (CONESCAPAN), San Salvador, El Salvador, 2022, pp. 1-6. DOI: 10.1109/CONESCAPAN56456.2022.9959592.

[14]D. Souto, et al., A New Type of Robot for Underwater Non-magnetic and Complex Hull Cleaning, Robotics and Automation, in: Proceedings of IEEE International Conference, Karlsruhe, Germany, 2013, pp. 3394-3399.
[15]V. Mazur, Robotic Mobile Platform for Container Transportation, in: 2021 IEEE 16th International Conference on the Experience of Designing and Application of CAD Systems (CADSM), Lviv, Ukraine, 2021, pp. 10-13. DOI: 10.1109/CADSM52681.2021.9385212.
[16]D. Canelon-Suarez, Y. Wang, N. Papanikolopoulos, Omnibot: A Small Versatile Robotic Platform Capable of Air, Ground, and Underwater Operation, in: 2020 International Conference on Unmanned Aircraft Systems (ICUAS), Athens, Greece, 2020, pp. 1742-1747. DOI: 10.1109/ICUAS48674.2020.9214021.
[17]A. Shevchenko, M. Vakulenko, M. Klymenko, The Ukrainian AI Strategy: Premises and Outlooks, in: Proceedings of the 12th Int. Conf. on Advanced Computer Information Technologies, ACIT 2022, Sep. 26-28, 2022, pp. 511-515. DOI: 10.1109/ACIT54803.2022.9913094
[18]R. Garcia-Hernandez et al., Decentralized Neural Control: Application to Robotics, Studies in Systems, Decision and Control, Vol. 96, Springer International Publishing, 2017.
[19]A.I. Shevchenko, M.S. Klymenko, Developing a Model of Artificial Conscience, in: 15th IEEE International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT'2020, vol. 1, Lviv-Zbarazh, 2020, pp. 51-54.
[20]M. Algabri, et al., Optimization of Fuzzy Logic Controller using PSO for Mobile Robot Navigation in an Unknown Environment, Applied Mechanics and Materials 541-542 (2014) 1053-1061.
[21]Y. P. Kondratenko, A. V. Kozlov, Parametric optimization of fuzzy control systems based on hybrid particle swarm algorithms with elite strategy, in Journal of Automation and Information Sciences 51, 12 (2019) 25-45. DOI: 10.1615/JAutomatInfScien.v51.i12.40
[22]H.A. Mayer, Ontogenetic teaching of mobile autonomous robots with dynamic neurocontrollers, International Journal of Computing 3, 1 (2004) 38-45.
[23]Y. -W. Song, J. Kang, S. -C. Yu, Development of Dual-Unit Ceiling Adhesion Robot System With Passive Hinge for Obstacle Traversal Under Kinodynamic Constraints, in: IEEE Access, vol. 11, 2023, pp. 4486-4500. DOI: 10.1109/ACCESS.2023.3236252.
[24]A. Papadimitriou, et al., On Adhesion Modeling and Control of a Vortex Actuator for Climbing Robots, in: 2019 IEEE 17th International Conference on Industrial Informatics (INDIN), Helsinki, Finland, 2019, pp. 571-576, DOI: 10.1109/INDIN41052.2019.8972244.
[25]A. Papadimitriou, G. Andrikopoulos, G. Nikolakopoulos, Experimental Evaluation of an Explicit Model Predictive Controller for an Adhesion Vortex Actuated Climbing Robot, in: 2020 American Control Conference (ACC), Denver, CO, USA, 2020, pp. 2137-2142. DOI: 10.23919/ACC45564.2020.9147658.
[26]S.-K. Kwon, et al., Re-adhesion control with estimated adhesion force coefficient for wheeled robot using fuzzy logic, 30th Annual Conference of IEEE Industrial Electronics Society, 2004, IECON 2004, Busan, Korea (South), Vol. 3, 2004, pp. 2530-2535. DOI: 10.1109/IECON.2004.1432200.
[27]Y. P. Kondratenko, et al., Automation of control processes in specialized pyrolysis complexes based on Web SCADA Systems, in: Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), in: Proceedings of the 9th IEEE International Conference. Bucharest, Romania, volume 1, 2017, pp. 107-112, DOI: 10.1109/IDAACS.2017.8095059.
[28]R. Jabłoński, T. Březina, (Eds.), Mechatronics, Recent Technological and Scientific Advances, Springer-Verlag Berlin Heidelberg, 2012.

[29] A. Trunov, et al., Functioning Model of the Ground Robotic Complex, in: International Scientific and Technical Conference on Computer Sciences and Information Technologies, Vol. 2, 2021, pp. 128-131.
[30] V.M. Kuntsevich, et al. (Eds.), Control Systems: Theory and Applications, River Publishers, Gistrup, Delft, 2018.
[31] C.C. Aggarwal, Artificial Intelligence, Springer International Publishing, 2021. DOI: 10.1007/978-3-030-72357-6.
[32] K. Knight, C. Zhang, G. Holmes, M.-L. Zhang, Artificial Intelligence, in: Proceedings of Second CCF International Conference, ICAI 2019, Xuzhou, China, August 22-23, 2019, Springer Singapore. DOI: 10.1007/978-981-32-9298-7.
[33] L.A. Zadeh, A.M. Abbasov, R.R. Yager, S.N. Shahbazova, M.Z. Reformat (Eds.), Recent Developments and New Directions in Soft Computing, STUDFUZ 317, Cham: Springer, 2014. DOI: 10.1007/978-3-319-06323-2.
[34] Y. P. Kondratenko, A. V. Kozlov, Generation of Rule Bases of Fuzzy Systems Based on Modified Ant Colony Algorithms, Journal of Automation and Information Sciences, 51, 3 (2019) 4-25. DOI: 10.1615/JAutomatInfScien.v51.i3.20.
[35] J. M. Mendel, Uncertain Rule-Based Fuzzy Systems, Introduction and New Directions, Second Edition, Springer International Publishing, 2017.
[36] Y.P. Kondratenko, O.V. Kozlov, Mathematical Model of Ecopyrogenesis Reactor with Fuzzy Parametrical Identification, in: Lotfi A. Zadeh et al. (Eds.), Recent Developments and New Direction in Soft-Computing Foundations and Applications, Studies in Fuzziness and Soft Computing, Vol. 342, Berlin, Heidelberg: Springer-Verlag, 2016, pp. 439-451. DOI: 10.1007/978-3-319-32229-2_30.
[37] D. Wang, et al., Neural optimal tracking control of constrained nonaffine systems with a wastewater treatment application, Neural Networks 143 (2021) 121-132.
[38] M. Lungu, R. Lungu, Neural network based adaptive control of airplane's lateral-directional motion during final approach phase of landing. Engineering Applications of Artificial Intelligence 74 (2018) 322-335.
[39] O. Kozlov, et al., Synthesis and Optimization of Green Fuzzy Controllers for the Reactors of the Specialized Pyrolysis Plants, in: Kharchenko V., Kondratenko Y., Kacprzyk J. (eds) Green IT Engineering: Social, Business and Industrial Applications, Studies in Systems, Decision and Control, Vol 171, Springer, Cham, 2019, pp. 373-396. DOI: 10.1007/978-3-030-00253-4_16
[40] E. Ferreyra et al., Depicting Decision-Making: A Type-2 Fuzzy Logic Based Explainable Artificial Intelligence System for Goal-Driven Simulation in the Workforce Allocation Domain, in: 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), New Orleans, LA, USA, 2019, pp. 1-6. DOI: 10.1109/FUZZ-IEEE.2019.8858933.
[41] M. Jamshidi, V. Kreinovich, J. Kacprzyk (Eds.), Advance Trends in Soft Computing, STUDFUZ 312, Cham: Springer-Verlag, 2013. DOI: 10.1007/978-3-319-03674-8.
[42] O. Castillo, P. Melin, An Approach for Optimization of Intuitionistic and Type-2 Fuzzy Systems in Pattern Recognition Applications, in: 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), New Orleans, LA, USA, 2019, pp. 1-5. DOI: 10.1109/FUZZ IEEE.2019.8858951.
[43] W. A. Lodwick, J. Kacprzhyk, Fuzzy Optimization, STUDFUZ, 254, Berlin, Heidelberg: Springer-Verlag, 2010. DOI: 10.1007/978-3-642-13935-2.
[44] O.V. Kozlov, Optimal Selection of Membership Functions Types for Fuzzy Control and Decision-Making Systems, in: Proceedings of the 2nd International Workshop on Intelligent Information Technologies & Systems of Information Security with CEUR-WS, Khmelnytskyi, Ukraine, IntelITSIS 2021, CEUR-WS, Vol-2853, 2021, pp. 238-247.

[45] B. Ullah, et al., Speed Control of Separately Excited DC Motor Using NARMA-L2 Controller, 2022 International Conference on Technology and Policy in Energy and Electric Power (ICT-PEP), Jakarta, Indonesia, 2022, pp. 43-47. DOI: 10.1109/ICT-PEP57242.2022.9988795.
[46] A. Nabi, N.A. Singh, Particle swarm optimization of fuzzy logic controller for voltage sag improvement. Proceedings of 2016 3rd International Conference on Advanced Computing and Communication Systems (ICACCS), Vol. 01, 2016, pp. 1-5.
[47] M.A. Elaziz, S. Mirjalili, A hyper-heuristic for improving the initial population of whale optimization algorithm. Knowledge-Based Systems, 172, 2019, pp. 42-63.
[48] A.-F. Attia, R.A. El Sehiemy, H.M. Hasanien, Optimal power flow solution in power systems using a novel Sine Cosine algorithm. International Journal of Electrical Power & Energy Systems 99 (2018) 331-343.
[49] H. Faris, et al. An efficient binary salp swarm algorithm with crossover scheme for feature selection problems. Knowledge-Based Systems 154 (2018) 43-67.
[50] S. Muthukaruppan, M.J. Er, A hybrid particle swarm optimization based fuzzy expert system for the diagnosis of coronary artery disease. Expert Systems with Applications, 39 (14), (2012) 11657-11665.
[51] J. Radosavljević, Optimal power flow using a hybrid optimization algorithm of particle swarm optimization and gravitational search algorithm. Electric Power Components and Systems, 43(17) (2015) 1958-1970.

УДК 681.51:004.896

НЕЧІТКА АВТОМАТИЧНА СИСТЕМА КЕРУВАННЯ АДГЕЗІЄЮ МОБІЛЬНОЇ РОБОТИЗОВАНОЇ ПЛАТФОРМИ

Зенг Ю.[1], Dr.Sci. Козлов О.[2], Донг С.[1], Ph.D. Кондратенко Г.[2], Скакодуб О.[2]

[1]*Політехнічний коледж Яньчен, Китай,*
E-mail: and_bsb@126.com, dongchengjian1988@126.com
[2] *Чорноморський національний університет імені Петра Могили, Україна,*
E-mail: kozlov_ov@ukr.net, halyna.kondratenko@chmnu.edu.ua, aleksandrskakodub1996@gmail.com

Анотація. *Розділ присвячений сучасним аспектам мобільної робототехніки, а саме універсальним мобільним роботизованим платформам, які можуть бути використані в різних технологічних середовищах і умовах для виконання різноманітних технічних завдань на промислових об'єктах і підприємствах. Ці мобільні платформи здатні пересуватися по горизонтальних поверхнях зі складним рельєфом, а також вертикально підніматися по стрімких стінах і стелях і є універсальним автономним засобом для виконання складних операцій у важкодоступних і небезпечних для людини місцях. Однією з найбільших проблем при експлуатації цих роботизованих платформ є належний контроль зчеплення під час руху по похилих і вертикальних площинах. У даній роботі авторами виконано проектування та дослідження нечіткої системи автоматичного керування зчепленням, яка забезпечує ефективне та надійне кріплення та переміщення платформи на похилих поверхнях різного типу. Запропонована система базується на механізмі нечіткого логічного висновку типу Мамдані, який дозволяє визначати необхідну силу зчеплення роботизованої платформи, що реалізує експертні знання для її безпечного та ефективного використання під різними кутами нахилу поверхні. Ефективність системи перевіряється комп'ютерним моделюванням.*

Ключові слова: *мобільна робототехніка, універсальна мобільна робототехнічна платформа, інтелектуальне керування адгезією, система нечіткого автоматичного керування, комп'ютерне моделювання.*

Section 2. Information control systems

UDC 681.391

EXTENDING NEURAL NETWORK MODELS WITH ADAPTIVE ACTIVATION FUNCTIONS

Kostiuk S.

Kharkiv National University of Radio Electronics, Ukraine
ORCID: 0000-0003-4196-2524
E-mail: serhii.kostiuk@nure.ua

Abstract. *Artificial neural network models continue to evolve in response to the new data processing tasks and challenges. To address new data processing requirements, researchers tend to develop new network architectures, update the existing ones, and seek abilities to improve the approximation capabilities for existing models. To speed up the innovation, the industry leans towards re-using the previously trained weights using the transfer learning techniques.This paper introduces the method of extending pre-trained artificial neural network models with adaptive activation functions. The method enables the usage of adaptive functions in place of the non-adaptive ones without the need for re-training. The practical aspects of the method and its performance are evaluated on the image classification task using a convolutional neural network model on an open source colour image dataset.*
Keywords: *adaptive activation function, neural network model, transfer learning, deep neural network*

1. Introduction

Artificial neural network models have become essential to modern data processing systems [1]. Convolutional neural networks (CNNs) serve as a core of image processing systems, enabling image processing at scale [2] and machine-assisted driving [3]. Recurrent neural networks (RNNs) can process sequential data, including video feeds [4] and historical data from industrial machinery [5]. The attention-based models, including the Transformer-based models, show state-of-the-art results in natural language translation [6], code generation [7], and language modeling [8-10]. Activation functions provide non-linearity to the network models. While the models in academic studies often use continuous activation functions (Tanh, Sigmoid, etc.), the commercial models tend to employ piece-wise linear activation functions (ReLU, PReLU, Hard Sigmoid, Hard Tanh, etc.) due to their lower computational complexity. At the same time, researchers continue developing new activation functions for different neural network architectures and data processing tasks [11]. One area of research is related to adaptive activation functions for deep neural networks. While networks with adaptive activation functions tend to perform better than the base variants [12], such networks are usually trained from scratch, limiting their applicability in production. To achieve better performance on the selected datasets, researchers tend to increase the number of trainable parameters in the model, add various layers to the network, and experiment with the overall structure of the neural network. An increase in the number of layers and parameters of the model leads to computational complexities in training and inference. As a result, there is a demand for optimizing the training process and using pre-trained models for new data processing tasks [13]. We introduce the method of extending pre-trained artificial neural network models with adaptive activation functions. We demonstrate the method using a CNN model implemented in PyTorch [14]. We compare the performance of the derived network, with and without activation function fine-tuning, to the performance of the base network on the CIFAR-10 dataset [15].

2. Related Works

Several research areas are related to the topic studied in this paper: non-adaptive activation functions development, adaptive activation functions synthesis, network architecture synthesis and manipulation, neural network pre-training, transfer learning, and fine-tuning. Works [11, 12, 13, 14, 15, 16] provide an overview of modern activation functions. Work [16] focuses on adaptive activation functions and describes the impact of adaptive activation functions on the effectiveness of the resulting model. The work studies parameterized standard and ensemble-based functions. The work highlights that the extensive comparison of adaptive activation functions requires their evaluation on the same dataset in the same base neural network architecture. However, the researchers cite results from other papers instead of conducting an independent assessment. Work [11] classifies activation functions by their characteristics, activation function properties, and target applications. As a piece-wise linear activation, ReLU benefits from simpler gradients and lower computational complexity. At the same time, ReLU shows moderately good results in the image processing and classification domains, leading to its common usage in commercial models. In contrast, the Sigmoid and Tanh activations are suboptimal for convolutional networks. The work studies Swish, ESwish, PReLU, and APL among the adaptive functions. As stated, the impact of additional parameters in adaptive activation functions on computational complexity is negligible compared to the overall number of parameters in a deep neural network. The authors compare the performance of models with different activation functions. The general model architecture remains the same across experiments, with activation functions being the only difference. The authors train the derived models from scratch and evaluate the performance on the same data set. The work highlights that the convergence of adaptive functions dramatically depends on their parameter initialization. In general, the models with adaptive functions demonstrate better convergence than their non-adaptive alternatives. Works [17, 18, 19, 20, 21] introduce new adaptive activation functions for deep neural networks, include them in different neural network models, and study their effectiveness on various datasets. The authors in [17] combine elementwise attention with the ReLU activation function to create a learnable activation function. The researchers in [20] replace all fixed activation functions in a BERT-based Transformer architecture with an adaptive Rational Activation Function (RAF). When trained from scratch, the RAF networks perform better than GELU on language modeling tasks. Work [19] introduces a continuous adaptive alternative to the ReLU and SiLU activation functions and demonstrates its ability to change the form and amplitude during training. Work [20] presents a piece-wise adaptive fuzzy activation function. Depending on its parameters and the number of membership functions, this fuzzy activation function can serve as a piece-wise approximation of continuous activations, such as Tanh and Sigmoid, on a pre-defined interval. Work [21] introduces a universal adaptive activation function and studies its ability to replace other non-adaptive activation functions. We note that works [17, 18, 19, 20, 21] only evaluate the effectiveness of models trained from scratch. The authors do not mention activation function replacement in pre-trained networks and do not reuse pre-trained parameters from non-adaptive implementations for further training. Works [22, 23, 24] propose and describe different artificial neural network models for text processing, including the CNN, RNN, and Transformer variants. In [22], the authors propose variants of Transformer and ImageNet with Swish as an adaptive sigmoid-based activation function. As stated, the models with Swish "show strong performance" compared to non-adaptive activations on the language translation and image classification tasks. [23] Introduces a deep CNN-based network for text sentiment analysis. The authors in [23] highlight the importance of unsupervised pre-training and learned embeddings for initializing deep text processing models. Work [24] presents DeepMoji, an RNN-based model for sentiment analysis. While [24] describes the architecture and the principle of operation of DeepMoji, it is only available in a pre-trained form without the original data set. Hence, independent researchers cannot modify the model, train it from scratch, and replicate the results from the paper using the same data set. Works [25, 26, 27, 28] study architecture manipulation for pre-trained neural network models. [25] overviews model compression methods for efficient deployment to resource-constrained devices and specialized hardware accelerators. The authors of [26] review the knowledge distillation methods. The reviewed methods use the teacher-student approach when the original (teacher) model remains unchanged while the teacher trains the new (student) model from scratch. In [27], the authors combine knowledge distillation with weight pruning for model compression. The authors apply fine-tuning to recover the model performance after pruning. Work [28] proposes the removal of weights close to zero and the associated connections between the nodes after each training epoch. To summarize, while works [25, 26, 27, 28] study modifications in the neural network architecture, they do not study the activation function replacement in such architectures.Works [13, 29, 30, 31, 31] study the application of fine-tuning, pre-training, and transfer learning techniques in data processing tasks. [13] overviews techniques for efficiently training large artificial neural network models, including pre-training and knowledge transfer. Some methods from this overview propose using pre-trained shallow models as the first layers of larger language modeling models. Other approaches include self-supervised and unsupervised pre-training of the whole model on a large dataset with fine-tuning on a target dataset. Overall, the authors highlight the importance and perspectives of pre-training for deep models like Transformers. Work [29] describes a multi-target evolutionary architecture search approach for CNN with weights sharing and sub-sampling of the target models. The trained network structure consists of pre-defined blocks, with a machine-readable configuration file as the structure definition. The authors apply fine-tuning on the target dataset to recover the accuracy of sub-sampled models. According to the paper and the reference implementation, the activation function remains fixed during the architecture search and training. Work [30] studies the contribution of individual layers in a recurrent Neural Machine Translation (NMT) model. The authors train a complete model on the bi-lingual out-of-domain OpenSubtitles2018 dataset, partially freeze the layers, and continue training on the target COPPA-V2 dataset. The experiment shows that the impact of individual layers during training is relatively small, as the model can still adapt to the new dataset with partially frozen layers. Work [31] overviews and evaluates the performance of transfer learning methods. To our knowledge, the previous works do not study the methods of activation function replacement in pre-trained artificial neural network models and the effectiveness of such replacement.

3. Method Design

Let N be an n-layer artificial neural network model with sequentially connected layers. A combination of its layers L serves as a network description: $N = \{L_1, L_2, \dots, L_n\}$. Each individual layer L_i implements a linear transformation together with a piece-wise non-linear transformation, implemented by activation function f_i: $L_i = f_i\left(X_i W_i^T + B_i\right)$, where X_i – the input tensor, W_i – trainable weights for linear transformation, B_i – the bias tensor, and f_i – the activation function. Let N_{orig} be the original pre-trained n-layer artificial neural network model, while N_{drv} – the derived network with its activation functions replaced. Hence N_{orig} and N_{drv} can be described as $N_{orig} = \{L_{orig,1}, L_{orig,2}, \dots, L_{orig,n}\}$ and $N_{drv} = \{L_{drv,1}, L_{drv,2}, \dots, L_{drv,n}\}$ correspondingly.

Depending on the implementation, networks N_{orig} and N_{drv} may share the same layers and activation functions for a subset of layers:

$$L_{orig,i} = L_{drv,i} = f_i\left(X_i W_i^T + B_i\right) \forall L_i \in L_{same}, L_{same} \subset L_{orig} \cap L_{drv}$$

.

As an alternative, the derived model may have all activation functions replaced while sharing the same values of trainable weights:

$$L_{orig,i} = f_{orig,i}\left(X_i W_i^T + B_i\right), L_{drv,i} = f_{drv,i}\left(X_i W_i^T + B_i\right)$$
$$\forall L_{i,orig} \in L_{orig}, L_{i,drv} \in L_{drv}, L_{orig} \cap L_{drv} = \emptyset.$$

Assuming the neural network architecture is known, the activation functions $f_{orig,i}$ can be replaced with the corresponding adaptive alternatives $f_{drv,i}$. As the output of each new layer L_i depends on the output of the previous layer L_{i-1}, changes in the hidden layers may change the distribution of their output values, introduce errors to their output values, invalidate the trainable parameter values in each subsequent layer, and require re-training of the following layers. The replacement activation function f_{drv} shall strictly follow the shape and amplitude of the original activation function f_{orig} to preserve the pre-trained parameter values W_i and B_i in the derived network. In other words, the difference between the outputs of the original and the replacement activation functions shall be zero for all input values z:

$$E_{f_i} = \left(f_{drv,i}(z) - f_{orig,i}(z)\right)^2 = 0 \ \forall \, z \in \mathbb{R}\,.$$

In practice, we show that the original activation function can be replaced by its approximation, assuming that the difference between their outputs on the whole range of input values is relatively low:

$$E_{f_i} = \left(f_{drv,i}(z) - f_{orig,i}(z)\right)^2 < \varepsilon \ \forall \, z \in \mathbb{R}\,.$$

3.1. Selection of Replacement Activation Functions

Achieving the zero difference between the outputs of the original and the replacement activation function is possible if the original activation function is a corner case of the replacement activation function. In this case, expressing the original activation function f_{orig} from the replacement adaptive activation function f_{drv} requires selecting the corresponding activation function parameters $P_{drv} = \{p_1; p_2; \dots; p_m\}$ so that:

$$f_{drv}(P_{drv}; z) = f_{orig}(z).$$

For example, Adaptive Hybrid Activation Function (AHAF) can exactly replace SiLU when $P_{drv} == \{p_1; p_2\} = \{1; 1\}$:

$$f_{orig}(z) = \text{SiLU}(z) = z\sigma(z),$$

$$f_{drv}(P_{drv}; z) = \text{AHAF}(p_1; p_2; z) = p_1 z\sigma(p_2 z),$$

$$f_{drv}(P_{drv}; z) = \text{AHAF}(1; 1; z) = 1 \cdot z\sigma(1 \cdot z) = z\sigma(z) = \text{SiLU}(z)$$

When parameter p_2 is close to $+\infty$, the sigmoidal part of AHAF gets close to the step function:

$$\lim_{p_2 \to +\infty} \sigma(p_2 z) = \sigma(+\infty \cdot z) = \begin{cases} 1, & z > 0 \\ 0, & z \le 0 \end{cases}$$

so that AHAF can approximate ReLU when $P_{drv} = \{p_1; p_2\} = \{1; +\infty\}$:

$$f_{orig}(z) = \text{ReLU}(z) = \begin{cases} z, & z > 0 \\ 0, & z \le 0 \end{cases}$$

$$f_{drv}(P_{drv}; z) = \text{AHAF}(1; +\infty; z) = 1 \cdot z\sigma(+\infty \cdot z) = \begin{cases} z, & z > 0 \\ 0, & z \le 0 \end{cases} = \text{ReLU}(z)$$

Modern computers have limited computation accuracy, so, in practice p_2 can be selected as a sufficiently large finite real number. Here and below, we use AHAF with $P_{drv} = \{p_1, p_2\} = \{1, 2^{16}\}$ as a sufficiently close approximation of ReLU:

$$f_{drv}(P_{drv}, z) = AHAF(1; 1 \cdot 2^{16}; z) = 1 \cdot z\sigma(2^{16} \cdot z) \approx \begin{cases} z, & z > 0 \\ 0, & z \le 0 \end{cases} \approx ReLU(z).$$

The same approach applies to replacing other adaptive activation functions, assuming that the replacement activation function is a generalization of the original one. For example, AHAF can replace the Swish activation function, when $P_{orig} = \{p_{orig,1}\}$, $p_{drv,1} = 1$, $p_{drv,2} = p_{orig,1}$, $P_{drv} = \{p_{drv,1}; p_{drv,2}\} = \{1; p_{orig,1}\}$:

$$f_{orig}(P_{orig}; z) = \text{Swish}(p_{orig,1}; z) = z\sigma(p_{orig,1} \cdot z),$$

$f_{drv}(P_{drv};z) = \text{AHAF}(p_{drv,1};p_{drv,2};z) = p_{drv,1} \cdot Z\sigma(p_{drv,2} \cdot z)$,
$f_{drv}(P_{drv};z) = \text{AHAF}(1;p_{orig,1};z) = 1 \cdot z\sigma(p_{orig,1} \cdot z) =$
$\text{Swish}(p_{orig,1};z)$

In contrast, some adaptive activation functions are piece-wise linear activation functions. Such adaptive activation functions can be an exact replacement only for a subset of other piece-wise linear activation functions. For example, the F-neuron activation function can replace the HardTanh activation function, assuming they have the same definition interval $[z_{min}, z_{max}]$:

$$f_{orig}(z) = \text{HardTanh}(z) = \begin{cases} z_{min}, & z < z_{min} \\ z_{max}, & z > z_{max} \\ z, & otherwise \end{cases},$$

$$f_{drv}(P,z) = \text{FN}_{\text{act}}(P,z) = \sum_{j=1}^{m}(FN_{act,j}(z) \cdot p_j) = \begin{cases} z_{min}, & z < z_{min} \\ z_{max}, & z > z_{max} \\ z, & otherwise \end{cases},$$

where $\text{FN}_{\text{act}}(P,z)$ is the F-neuron activation function, $\text{FN}_{\text{act},j}(z)$ is the j-th membership function of $\text{FN}_{\text{act}}(P,z)$, $P = \{p_1, p_2, \dots, p_m\}$ – parameters of the F-neuron activation function, m – the total number of membership functions,

$p_j = \text{HardTanh}(z_{min} + (z_{max} - z_{min})/(m-1) * (j-1))$.

When the original function is not piece-wise linear, the F-neuron activation function can only provide a piece-wise linear approximation on an interval with limited accuracy. For example, an F-neuron with 18 membership functions (1 left ramp, 16 triangle-shaped functions, 1 right ramp) can approximate Tanh on the $[-4.0; +4.0]$ interval with the maximum delta squared of $4 \cdot 10^{-4}$. Figure 1 illustrates the delta squared error for the described F-neuron.In this case, the trainable parameters $P_{FN_{act}}$ are initialized as:

$$P_{FN_{act}} = \{p_1, p_2, \dots, p_{18}\}, \text{ m} = 18,$$

$$p_j = \tanh\left(-4.0 + \frac{4.0+4.0}{17} * (j-1)\right),$$

$$P_{FN_{act}} \approx \{-0.9993; -0.9992; \ldots; -0.031; +0.031; \ldots; \ldots; 0.9992; 0.9993\}$$

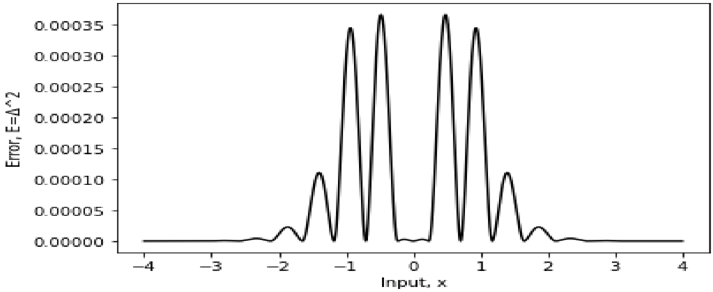


Fig. 1. Delta squared error for an F-neuron activation initialized as Tanh with 18 membership functions compared to the base Tanh activation function.

3.2. Practical Aspects of Activation Function Replacements

Researchers and companies use various data formats to describe, save and distribute their pre-trained neural network models. To improve interoperability and simplify the practical application of the model, the authors use various deep learning frameworks, such as PyTorch [14] and others. Generally, such frameworks store their models in two separate pieces: the model description in code and the values of its trainable weights in the binary files. Hence, the implementation of the activation function replacement method is a five-step procedure:

1. Restore the original neural network structure in memory using the stored model description.

2. Replace the original activation functions with their adaptive alternatives to get the derived network structure.

Load values of the original trainable parameters from the saved state.
Initialize the parameters of the adaptive functions so that the adaptive functions correspond to the original ones, as described in Section 3.1. (Optional) Fine-tune the replacement activations.

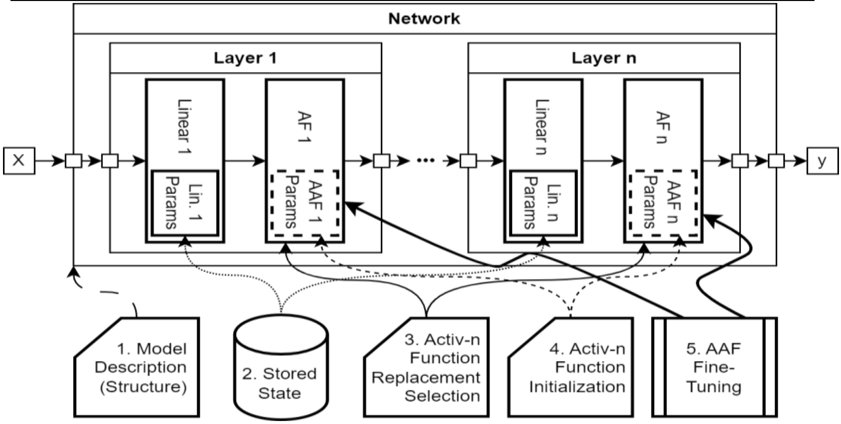


Fig. 2. Network model initialization and activation function replacement from the description of the original model description, the stored state, and the replacement rules described in Section 3.1.

Figure 2 illustrates the network initialization and activation function replacement procedure.

In practice, modern deep learning frameworks allow combining steps 1, 3, and 4. The code sample in Figure 3 shows an example two-step implementation in PyTorch.

```
1    net = NetworkWithAdaptiveActivations()
     saved_state = torch.load(saved_state_file_path)
2    net.load_state_dict(saved_state, strict=False)
```

Fig. 3. Loading pre-trained weights into a derived network with adaptive activation functions.

In this figure, “net” indicates a new network in memory, “saved_state_file_path” denotes the pre-trained model’s state path, and “strict=False” instructs PyTorch to load parameters of the pre-trained network, ignoring the adaptive activation function parameters missing from the original model.

The fine-tuning step can be implemented as the general training procedure with all synaptic weights frozen, except the ones belonging to adaptive activation functions.

3.3. Fine-Tuning Activation Function Parameters in a Pre-Trained Model

An imperfect approximation of the original activation function may cause degradation in the artificial neural network model's performance (classification accuracy, loss function values, etc.). The replacement activation functions may provide an imperfect approximation of the original ones. The adaptive activation function parameters can be fine-tuned separately from all other trainable parameters to improve the performance and compensate for possible approximation errors. The non-activation parameters shall be excluded from the training during the fine-tuning process. The code sample in Figure 4 illustrates the approach: Where "net" is an instance of the artificial neural network model, "dataset" is the training data set, "net.params()" returns all trainable parameters in the network, and "net.activation_params" returns the list of parameters that belong to adaptive activation functions.

```
for p in net.params():
    p.requires_grad = False

for p in net.activation_params:
    p.requires_grad = True

train_network(net, dataset)
```

Fig. 4. Fine-tuning the adaptive activation function parameters.

In the following sections, we empirically study the effectiveness of networks with fine-tuned adaptive activation functions compared to the original networks and networks with adaptive activation functions that were trained from scratch.

4. Experiment

We evaluate the performance of various neural network variants on the CIFAR-10 [15] data set. Each element in the data set is a 3-channel image, 32x32 pixels in size. Each image in the data set corresponds to one of 10 possible object classes. We use the 5:1 split between the training and the testing data: 50000 images in the training subset and 10000 in the testing subset. During training, we apply image augmentation using a random horizontal flip and a random affine transformation (shifting) by at most 0.1 times vertically and horizontally. The base neural network architecture is the VGG-like KerasNet [32] network. The architecture includes: 4 two-dimensional convolutional layers with the corresponding activation functions; 2 layers of 2D max pooling and 2D dropout between the pairs of convolutional layers; 1 hidden fully connected layer with the corresponding activation function; 1 layer of dropout between the fully connected layers; 1 output fully connected layer with SoftMax on its output for classification.

The architecture changes based on the model variant (the reference network, the network with AHAF, the network with the F-neuron activation). Still, the overall list of layers and the list of non-activation components remains the same across all experiments. Unless otherwise noted, we run the experiments on a laptop with Intel Core i7 10510U, NVIDIA GeForce GTX 1650 Max-Q, and PyTorch [14] version 1.13.1. During the training and evaluation, we use the default floating point data type for this version of PyTorch – float32. The implementation is available on GitHub: https://github.com/s-kostyuk/better_nns_w_aafs. Compared to the previous studies [33], we have updated the implementation, re-run the experiments in the new reproducible environment, and collected additional data to investigate the reliability of the obtained results, prove the correctness of the proposed methods and implementations in different training conditions.

4.1. Reference Model Training from Scratch

We train four variants of the reference KerasNet model: With the ReLU activation in all layers; With SiLU in all layers; With Tanh in all layers; - With Sigmoid in all layers. We train each variant from scratch for one hundred epochs on the CIFAR-10 dataset. We record the test set accuracy and the training set loss to visualize and analyze the training process. We save the resulting weights for further reference and activation function replacement. This experiment provides the reference data required for comparison with the derived models.

4.2. Replacing Activation Functions in Pre-Trained Reference Model

We create the derivative models from the reference models by loading the pre-trained weights and replacing the non-adaptive activation functions with the adaptive ones. This step generates four different model variants:

1. With the AHAF activation in all layers, AHAF initialized as ReLU.
2. With the AHAF activation in all layers, AHAF initialized as SiLU.

3. With Tanh activation in the convolutional layers and the F-neuron activation in the second-to-last fully connected layer, the F-neuron activation approximates Tanh with 18 membership functions.

4. With Sigmoid activation in the convolutional layers and the F-neuron activation in the second-to-last fully connected layer, the F-neuron approximates Sigmoid with 18 membership functions.

We evaluate the performance of derived networks directly after the replacement. This experiment allows validating the replacement approach and studies the effect of imperfect replacements when used with piece-wise linear adaptive activation functions.

4.3. Activation Function Fine-Tuning

We fine-tune the activation function parameters in the derived models to compensate for approximation errors and improve the classification performance. We freeze all trainable parameters, except the activation function parameters, during fine-tuning. We fine-tune for 50 additional epochs on the same CIFAR-10 dataset, so there is no need to adapt the architecture for a new dataset. We record the test set accuracy and the training set loss during the fine-tuning process. We save the activation function parameters for subsequent visualization. This experiment evaluates the impact of activation function fine-tuning on the model performance and the form of activation functions.

4.4. Training Models with Adaptive Activation Functions from Scratch

We train four additional variants of the KerasNet model from scratch:

1. With the AHAF activation in all layers, AHAF initialized as ReLU.
2. With the AHAF activation in all layers, AHAF initialized as SiLU.
3. With Tanh activation in the convolutional layers and the F-neuron activation in the second-to-last fully connected layer, the F-neuron activation approximates Tanh with 18 membership functions.
4. With Sigmoid activation in the convolutional layers and the F-neuron activation in the second-to-last fully connected layer, the F-neuron approximates Sigmoid with 18 membership functions.

We record the training time, the test set accuracy, and the training set loss. This experiment compares the training time and effectiveness between models with replaced activation functions and those trained initially with adaptive activations.

4.5. Additional Experiments

To validate the results, we perform two additional sets of experiments. The first set investigates the accuracy of replacement activation functions for different floating-point data types, and the second studies the training results reproducibility with different seed values. Due to the nature of floating-point computations on discrete computers, the mathematically identical computations might lead to different results with some degree of accuracy. As the bit length and the method of the floating-point value storage changes, the original and the replacement activation functions show different degrees of similarity. The goal of this set of experiments is to demonstrate the impact of floating-point data types on the activation function replacement accuracy. In the second set of experiments, we repeat the previous experiments using different random seed values and, as the result, different starting values for the synaptic weights. The goal of this set of experiments is to validate and to isolate the noise in the previously obtained results. To speed up the process, we use a different environment with NVIDIA RTX A4000 and PyTorch version 2.0.

5. Results and Discussions

We recorded and evaluated the results of the experiments. We structure this section as the following: Validation of the parameter initialization and the form for adaptive activation functions. Evaluation of the activation function form after pre-training with all other parameters frozen. Comparison of the performance recordings for different variants of the KerasNet model.

5.1. Validation of Activation Function Replacement

We visualize the activation function form for adaptive activation functions directly after the activation function replacement. As the visualization shows, the form of AHAF activations (grey; AHAF-as-RELU in Figure 5, and AHAF-as-SiLU in Figure 6, correspondingly) directly follows the form of the corresponding original functions (black; ReLU and SiLU correspondingly). Hence, we confirm that the AHAF activation can directly replace ReLU and SiLU activations.

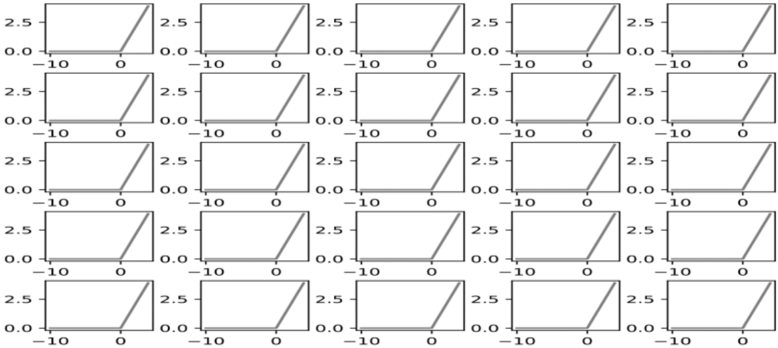


Fig. 5. The form of ReLU-like AHAF after replacement.

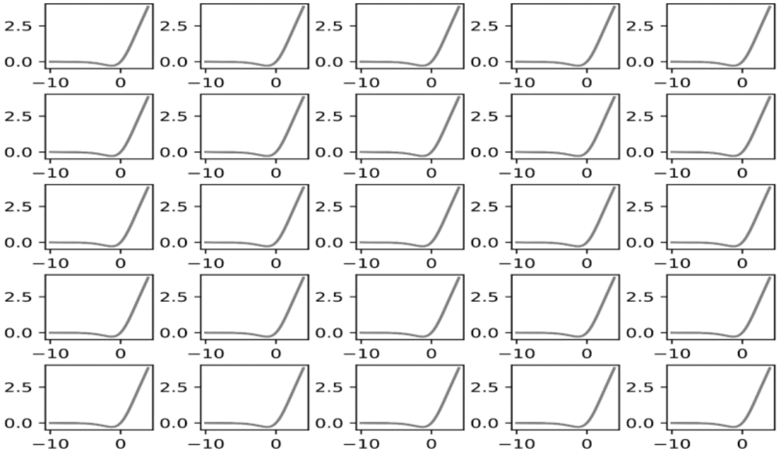


Fig. 6. The form of SiLU-like AHAF after replacement.

The visualizations for F-neuron activations (Figure 7 and Figure 8 for Tanh-like and Sigmoid-like F-neurons, correspondingly) show that while the replacement is not perfect, the differences between the original functions (black) and the replacement function (grey) are hard to spot. Hence, the parameter initialization is correct for the F-neuron activations.

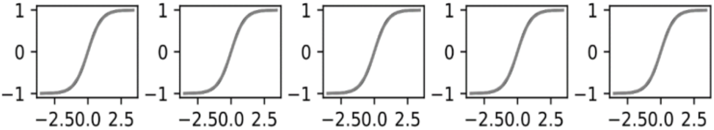


Fig. 7. Activation form of the F-neuron activation as Tanh after replacement

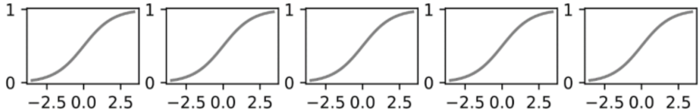


Fig. 8. Activation form of the F-neuron activation as Sigmoid after replacement

5.2. Adaptive Activations' Form After Fine-Tuning

Figures 9-12 visualize the activation function form for the AHAF and F-neuron activation variants after the fine-tuning. The visualization shows that the form of fine-tuned activation functions (grey) noticeably diverges from the original form (black). We note serious deformation of the sigmoid-like F-neuron activations which indicates poor applicability of Sigmoid for convolutional networks. Hence, we confirm that the activation function fine-tuning works and the activation function form changes depending on the requirements of the model.

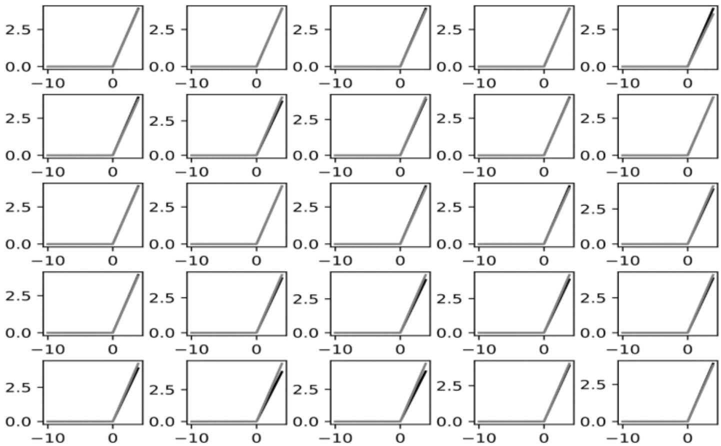


Fig. 9. Activation form of AHAF as ReLU after fine-tuning

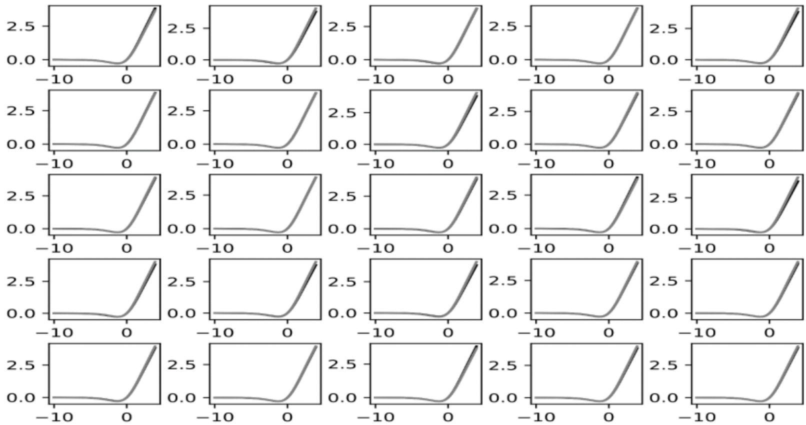


Fig. 10. Activation form of AHAF as SiLU after fine-tuning

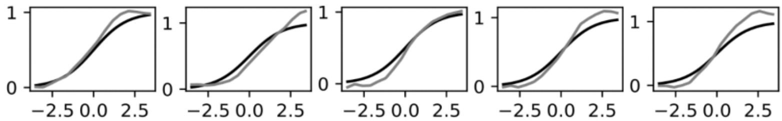


Fig. 11. Activation form of the F-neuron activation as Tanh after fine-tuning

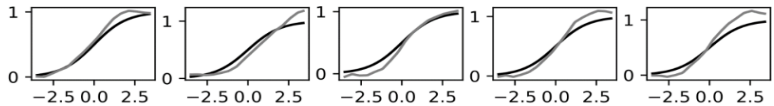


Fig. 12. Activation form of the F-neuron activation as Sigmoid after fine-tuning

5.3. Performance Of the Model Variants

We compare the test set accuracy between all variants of the KerasNet model. We group by the results by the reference activation functions used in such networks: Linear units – SiLU-like and ReLU-like; Bounded functions – Tanh-like and Sigmoid-like. The fine-tuned model with the ReLU-like AHAF activations shows the best test set accuracy on the CIFAR-10 dataset. AHAF successfully replaces the original ReLU activation functions and adapts its form during fine-tuning but tends to overfit on the final epochs. While the overall training time is longer, activation function replacement together with activation function fine-tuning shows itself as a viable solution when the pre-trained model with non-adaptive functions is already available.

As expected, the test set accuracy after the AHAF-based activation function replacement is the same as the accuracy of the reference models after one hundred epochs.

Table 1 demonstrates the best training results for models with linear unit activations.

The model with the Tanh-like F-neuron activation trained from scratch shows the best test set accuracy on the CIFAR-10 dataset. One potential explanation for such results is the poor applicability of Tanh and Sigmoid activations for convolutional neural networks.

Training the model from scratch allows modification of the activation function form and hence better propagation of the signal between the layers. As expected, the F-neuron-based activation function replacement gives different results to the reference model, but the actual difference is negligible (less than 0.01% for Tanh and about 0.12% for Sigmoid).

Even for networks with Tanh-like and Sigmoid-like activations, activation function replacement together with activation function fine-tuning is a viable option when the pre-trained model is already available.

Table 2 demonstrates the best training results for models with bounded activations.

Table 1
Performance of model variants with SiLU-like and ReLU-like activations.

Variant	CNN AF	FFN AF	Accuracy, %	Training Epoch	Avg Epoch Time, seconds
Reference	ReLU	ReLU	82.41	96	13.52
AHAF, from scratch	ReLU-like	ReLU-like	82.29	99	19.85
AHAF, AF replaced	ReLU-like	ReLU-like	82.17	100	-
AHAF, fine-tuned AF	ReLU-like	ReLU-like	**82.65**	148	14.66
Reference	SiLU	SiLU	81.59	98	15.65
AHAF, from scratch	SiLU-like	SiLU-like	81.85	98	19.82
AHAF, AF replaced	SiLU-like	SiLU-like	81.53	100	-
AHAF, fine-tuned AF	SiLU-like	SiLU-like	81.96	147	14.64

Table 2

Performance of model variants with Tanh-like and Sigmoid-like activations.

Variant	CNN AF	FFN AF	Accuracy, %	Training Epoch	Avg Epoch Time, seconds
Reference	Tanh	Tanh	79.97	100	13.05
Fuzzy, from scratch	Tanh	Tanh-like	**81.04**	100	13.30
Fuzzy, AF replaced	Tanh	Tanh-like	79.97	100	-
Fuzzy, fine-tuned AF	Tanh	Tanh-like	80.34	150	11.37
Reference	Sigmoid	Sigmoid	60.17	100	14.54
Fuzzy, from scratch	Sigmoid	Sigmoid-like	60.55	100	13.17
Fuzzy, AF replaced	Sigmoid	Sigmoid-like	60.43	100	-
Fuzzy, fine-tuned AF	Sigmoid	Sigmoid-like	61.16	147	11.66

Figures 13 and 14 demonstrate the training process for models with linear units and bounded functions correspondingly.

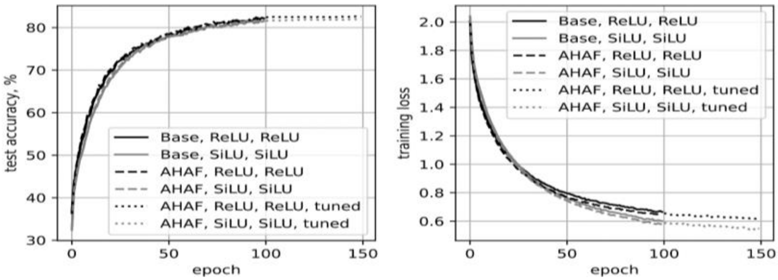


Fig. 13. The training process for network variants with SiLU and ReLU activations

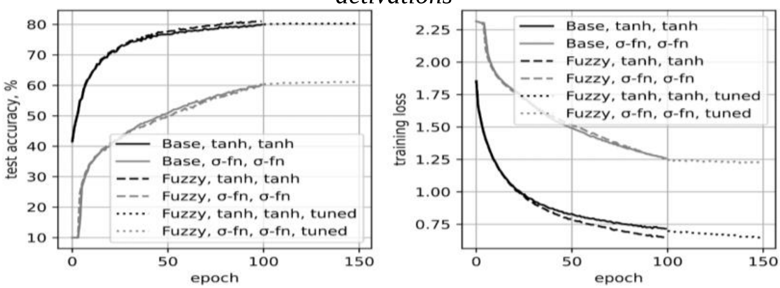


Fig. 14. The training process for network variants with Tanh and Sigmoid activations

5.4. Floating Point Impact on Activation Function Accuracy

The modern deep learning frameworks and computing devices use imperfect floating-point representations. The differences in the order of operations, the flowing point precision, and the implementation details of specific computing devices result in subtle variations in the loss value, gradients, and the network performance after the training. For deep neural networks, such errors accumulate and may lead to performance divergence between the two, seemingly identical, models. Such differences should be accounted for when comparing different neural network models. Experimentally, we observe differences between the built-in SiLU implementation in PyTorch (torch.nn.SiLU) and the mathematically equivalent AHAF implementation. The maximum delta squared error between the two implementations ranges from $3.16 \cdot 10^{-30}$ for the float64 data type to $6.1 \cdot 10^{-5}$ for the float16 data type. Figure 15 illustrates the delta squared error between SiLU implementations for different data types.

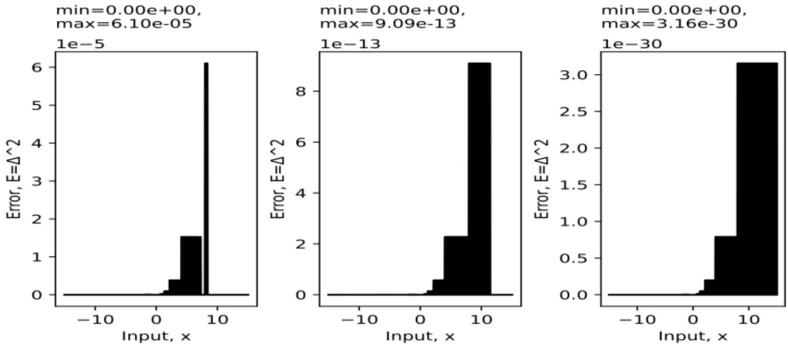


Fig. 15. The delta squared error between the native and the from-scratch SiLU implementations for float16 (left), float32 (center), and float64 (right)

For the piece-wise linear approximation, the approximation quality mainly depends on the number of linear pieces and the parameter initialization procedures. For example, the Tanh-like F-neuron activation instance with 18 membership functions and the parameter values sampled from Tanh, shows a relatively high delta squared error of $3.72 \cdot 10^{-4}$ for float16 and $3.66 \cdot 10^{-4}$ for higher precisions, as illustrated on Figure 16. This fact explains the accuracy differences between the base model and the derivative model with the F-neuron activations.Alternatively, the Tanh-like F-neuron activation instance with 130 membership functions and same parameter initialization procedures, shows delta squared error of $9.54 \cdot 10^{-7}$ for float16 and $3.36 \cdot 10^{-7}$ for higher precisions. While showing better results than the simpler alternative, the F-neuron activation with 130 membership functions shows error values that are outside of the noise range for floating point computations. Figure 17 illustrates the error for the described activation function. As shown, the error values grow as the input values move outside of the F-neuron activation definition range (-4.0 to +4.0). This hints on the potential for improvements in the F-neuron activation initialization procedures to reduce the error outside of the definition boundaries.

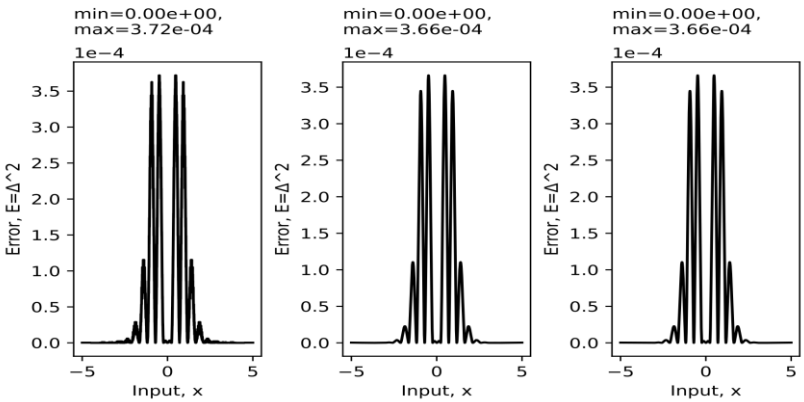


Fig. 16. The delta squared error between the native Tanh implementation and its fuzzy approximation with 18 membership functions for float16 (left), float32 (center), and float64 (right)

For AHAF-as-ReLU, the error values on GPU vary between zero for float16 and $1.21 \cdot 10^{-13}$ for float32. As the input values approach zero, the difference between the base ReLU function and its AHAF approximation grows, resulting from the exponential nature of AHAF. In theory, the approximation quality should increase together with increasing the p_2 parameter of AHAF. In practice, such an imperfect approximation is sufficient for practical applications. Figure 18 shows the error between the base ReLU activation and its AHAF-as-ReLU approximation with $\{p_1, p_2\} = \{1, 2^{16}\}$.

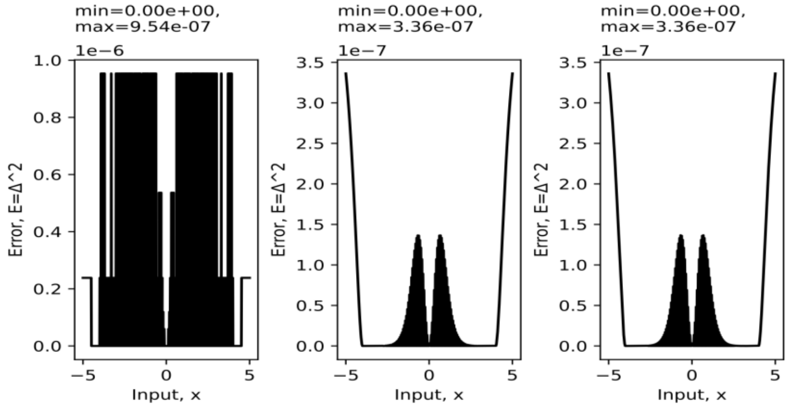


Fig. 17. The delta squared error between the native Tanh implementation and its fuzzy approximation with 130 membership functions for float16 (left), float32 (center), and float64 (right)

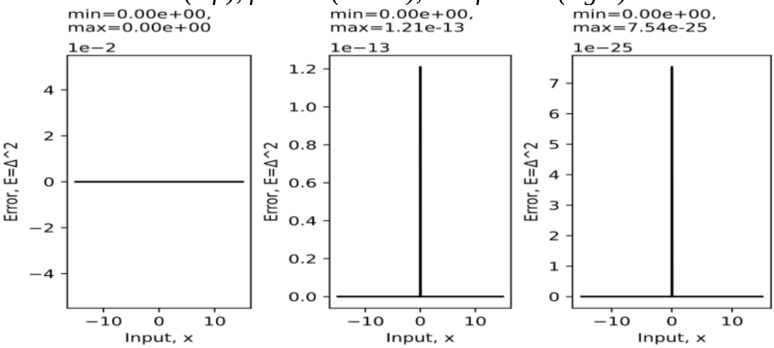


Fig. 18. The delta squared error between the native ReLU implementation and its AHAF approximation for float16 (left), float32 (center), and float64 (right) evaluated on a GPU

Interestingly, the error values differ between the CPU and GPU executors.

Compared to GPU, the error values for the float32 and float64 match exactly when evaluated on CPU, as illustrated by Figure 19.

The diagram does not include the float16 case as the CPU does not support this data format.

The full set of error graphs for ReLU, SiLU, and Tanh is available on a reasonable request.

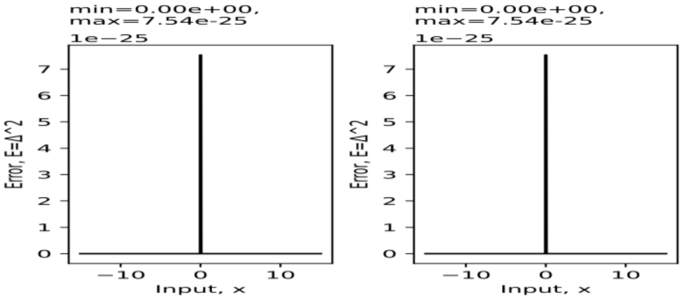


Fig. 19. The delta squared error between the native ReLU implementation and its AHAF approximation for float32 (left) and float64 (right) evaluated on a CPU

5.5. The Seed Value Impact on Reproducibility

We repeated the training experiments using 5 different seed values: 100, 128, 1999, 7823, and 42. Compared to the original set of experiments, we used a different computer with NVIDIA RTX A4000 and PyTorch version 2.0.

Table 3

Model performance variance across different starting seed values.

Variant	CNN AF	FFN AF	Test Set Loss, Mean	Test Set Loss, SD	Test Set Acc., Mean, %
Reference	ReLU	ReLU	0.5124	0.01587	81.76
AHAF, from scratch	ReLU-like	ReLU-like	0.5110	0.00648	81.88
AHAF, fine-tuned AF	ReLU-like	ReLU-like	0.4932	0.01756	82.34
Reference	SiLU	SiLU	0.5117	0.01042	81.75
AHAF, from scratch	SiLU-like	SiLU-like	0.5046	0.00829	81.90
AHAF, fine-tuned AF	SiLU-like	SiLU-like	0.5039	0.01036	82.12
Reference	Tanh	Tanh	0.5734	0.01365	79.77
Fuzzy, from scratch	Tanh	Tanh-like	0.5483	0.01233	80.60
Fuzzy, fine-tuned AF	Tanh	Tanh-like	0.5545	0.01183	80.25
Reference	Sigmoid	Sigmoid	1.1434	0.01917	59.74
Fuzzy, from scratch	Sigmoid	Sigmoid-like	1.1564	0.02340	59.64
Fuzzy, fine-tuned AF	Sigmoid	Sigmoid-like	1.1197	0.01466	60.89

This change allowed us to speed up the training and account for possible machine-to-machine and execution environment variations. Table 3 shows the mean and the standard deviation for the test set loss, and the mean test set accuracy value for each of the network variants at the end of the training – after 100 epochs for networks trained from scratch, after 50 fine-tuning epochs for fine-tuned epochs. According to the evaluation results, the test set accuracy and loss remain largely consistent regardless of the initial seed values. On average, activation function fine-tuning allows reaching the best test set accuracy for AHAF-as-ReLU, AHAF-as-SiLU, and F-neuron activation in the Sigmoid-like variant.

The difference between the corresponding SiLU and ReLU variants is insignificant and can be regarded as noise. Table 4 illustrates the variations between machines using the seed of 42.

The interactive presentation of all training runs is available at the Weights and Biases platform: https:/api.wandb.ai/links/nure-cslr-1/s5zq6yp4.

Table 4
Model performance variance across different machines with the constant seed of 42.

Variant	CNN AF	FFN AF	Test Set Loss, PC	Test Set Loss, laptop	Test Set Acc., PC	Test Set Acc., laptop
Reference	ReLU	ReLU	0.50369	0.50172	82.35%	82.17%
AHAF, from scratch	ReLU-like	ReLU-like	0.50133	0.51730	81.92%	82.07%
Reference	SiLU	SiLU	0.51276	0.51268	81.58%	81.53%
AHAF, from scratch	SiLU-like	SiLU-like	0.50215	0.50182	81.85%	81.84%
Reference	Tanh	Tanh	0.58671	0.58647	80.13%	79.97%
Fuzzy, from scratch	Tanh	Tanh-like	0.54351	0.54275	80.70%	81.04%
Reference	Sigmoid	Sigmoid	1.12710	1.12360	60.20%	60.17%
Fuzzy, from scratch	Sigmoid	Sigmoid-like	1.14425	1.12749	59.55%	60.55%

6. Conclusions

This paper introduces the method of extending pre-trained artificial neural network models with adaptive activation functions. The method allows experimenting with various adaptive activation functions without the need for model re-training. We evaluate several variants of the KerasNet model on the CIFAR-10 dataset: the reference pre-trained model with non-adaptive activation function, the derived model with replaced activation functions, and the derived model trained from scratch. The empirical results confirm the possibility of using piece-wise approximations of Sigmoid and Tanh without significant changes in the test set accuracy. The experiment shows the effectiveness of the method in combination with activation function fine-tuning for improving the accuracy of models.

7. References

[1] A. Paleyes, R.-G. Urma, N. D. Lawrence, Challenges in deploying machine learning: A survey of case studies, ACM Computing Surveys 55 (2022) 1–29. DOI:10.1145/3533378.

[2] E. Yurtsever, J. Lambert, A. Carballo, K. Takeda, A survey of autonomous driving: Common practices and emerging technologies, IEEE Access 8 (2020) 58443–58469. DOI:10.1109/access.2020.2983149.

[3] S. S. A. Zaidi, M. S. Ansari, A. Aslam, N. Kanwal, M. Asghar, B. Lee, A survey of modern deep learning based object detection models, Digital Signal Processing 126 (2022) 103514. DOI:10.1016/j.dsp.2022.103514.

[4] S. Grigorescu, B. Trasnea, T. Cocias, G. Macesanu, A survey of deep learning techniques for autonomous driving, Journal of Field Robotics 37 (2020) 362–386. DOI:10.1002/rob.21918.

[5] Y. Wang, M. Perry, D. Whitlock, J. W. Sutherland, Detecting anomalies in time series data from a manufacturing system using recurrent neural networks, Journal of Manufacturing Systems 62 (2022) 823–834. DOI:10.1016/j.jmsy.2020.12.007.

[6] S. Ranathunga, E.-S. A. Lee, M. P. Skenduli, R. Shekhar, M. Alam, R. Kaur, Neural machine translation for low-resource languages: A survey, ACM Computing Surveys 55 (2023) 1–37. DOI:10.1145/3567592.

[7] F. F. Xu, U. Alon, G. Neubig, V. J. Hellendoorn, A systematic evaluation of large language models of code, in: Proceedings of the 6th ACM SIGPLAN International Symposium on Machine Programming, ACM, 2022. DOI:10.1145/3520312.3534862.

[8] GPT-4 Technical Report, Technical Report, OpenAI, 2023. DOI:10.48550/ARXIV.2303.08774.

[9] T. Wolf, L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi, P. Cistac, T. Rault, R. Louf, M. Funtowicz, J. Davison, S. Shleifer, P. von Platen, C. Ma, Y. Jernite, J. Plu, C. Xu, T. L. Scao, S. Gugger, M. Drame, Q. Lhoest, A. Rush, Transformers: State-of-the-art natural language processing, in: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, Association for Computational Linguistics, 2020. DOI:10.18653/v1/2020.emnlp-demos.6.

[10] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, in: Proceedings of the 2019 Conference of the North, Association for Computational Linguistics, 2019. DOI:10.18653/v1/n19-1423.

[11] S. R. Dubey, S. K. Singh, B. B. Chaudhuri, Activation functions in deep learning: A comprehensive survey and benchmark, Neurocomputing 503 (2022) 92–108. DOI:10.1016/j.neucom.2022.06.111.

[12] A. D. Jagtap, K. Kawaguchi, G. E. Karniadakis, Adaptive activation functions accelerate convergence in deep and physics-informed neural networks, Journal of Computational Physics 404 (2020) 109136. DOI:10.1016/j.jcp.2019.109136.

[13] X. Han, Z. Zhang, N. Ding, Y. Gu, X. Liu, Y. Huo, J. Qiu, Y. Yao, A. Zhang, L. Zhang, W. Han, M. Huang, Q. Jin, Y. Lan, Y. Liu, Z. Liu, Z. Lu, X. Qiu, R. Song, J. Tang, J.-R. Wen,

J. Yuan, W. X. Zhao, J. Zhu, Pre-trained models: Past, present and future, AI Open 2 (2021) 225–250. DOI: 10.1016/j.aiopen.2021.08.002.

[14] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, Pytorch: An imperative style, high-performance deep learning library, in: H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, R. Garnett (Eds.), Advances in Neural Information Processing Systems, volume 32, Curran Associates, Inc., 2019.

[15] A. Krizhevsky, G. Hinton, Learning multiple layers of features from tiny images, Technical Report, University of Toronto, Toronto, Ontario, 2009.

[16] A. Apicella, F. Donnarumma, F. Isgrò, R. Prevete, A survey on modern trainable activation functions, Neural Networks 138 (2021) 14–32. DOI:10.1016/j.neunet.2021.01.026.

[17] D. Chen, K. Xu, Arelu: Attention-based rectified linear unit, CoRR abs/2006.13858 (2020).

[18] H. Fang, J.-U. Lee, N. S. Moosavi, I. Gurevych, Transformers with learnable activation functions, 2022. DOI:10.48550/ARXIV.2208.14111.

[19] Y. Bodyanskiy, S. Kostiuk, Adaptive hybrid activation function for deep neural networks, System research and information technologies (2022) 87–96. DOI:10.20535/srit.2308-8893.2022.1.07.

[20] Y. Bodyanskiy, S. Kostiuk, Deep neural network based on f-neurons and its learning (2022). DOI:10.21203/rs.3.rs-2032768/v1.

[21] B. Yuen, M. T. Hoang, X. Dong, T. Lu, Universal activation function for machine learning, Scientific Reports 11 (2021). DOI:10.1038/s41598-021-96723-8.

[22] P. Ramachandran, B. Zoph, Q. V. Le, Searching for activation functions, in: 6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Workshop Track Proceedings, OpenReview.net, 2018.

[23] A. Severyn, A. Moschitti, Twitter sentiment analysis with deep convolutional neural networks, in: Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, ACM, 2015. DOI:10.1145/2766462.2767830.

[24] B. Felbo, A. Mislove, A. Søgaard, I. Rahwan, S. Lehmann, Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm, in: Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, 2017. DOI:10.18653/v1/d17-1169.

[25] B. L. Deng, G. Li, S. Han, L. Shi, Y. Xie, Model compression and hardware acceleration for neural networks: A comprehensive survey, Proceedings of the IEEE 108 (2020) 485–532. DOI:10.1109/jproc.2020.2976475.

[26] J. Gou, B. Yu, S. J. Maybank, D. Tao, Knowledge distillation: A survey, International Journal of Computer Vision 129 (2021) 1789–1819. DOI:10.1007/s11263-021-01453-z.

[27] N. Aghli, E. Ribeiro, Combining weight pruning and knowledge distillation for CNN compression, in: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE, 2021. DOI:10.1109/cvprw53098.2021.00356.

[28] D. C. Mocanu, E. Mocanu, P. Stone, P. H. Nguyen, M. Gibescu, A. Liotta, Scalable training of artificial neural networks with adaptive sparse connectivity inspired by network science, Nature Communications 9 (2018). DOI:10. 1038/s41467-018-04316-3.

[29] Z. Lu, G. Sreekumar, E. Goodman, W. Banzhaf, K. Deb, V. N. Boddeti, Neural architecture transfer, IEEE Transactions on Pattern Analysis and Machine Intelligence 43 (2021) 2971–2989. DOI:10.1109/tpami.2021.3052758.

[30] B. Thompson, H. Khayrallah, A. Anastasopoulos, A. D. McCarthy, K. Duh, R. Marvin, P. McNamee, J. Gwinnup, T. Anderson, P. Koehn, Freezing subnetworks to analyze domain adaptation in neural machine translation, in: Proceedings of the Third Conference on Machine

Translation: Research Papers, Association for Computational Linguistics, 2018. DOI:10.18653/v1/w18-6313.
[31]F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, Q. He, A comprehensive survey on transfer learning, Proceedings of the IEEE 109 (2021) 43–76. DOI:10.1109/jproc.2020.3004555.
[32]F. Chollet, Train a simple deep cnn on the cifar10 small images dataset — keras 1.2.2., 2017.
[33]Y. Bodyanskiy, S. Kostiuk, Using Adaptive Activation Functions in Pre-Trained Artificial Neural Network Models, in: materials of the XI international scientific-practical conference "Information Control Systems and Technologies (ICST-Odessa-2023)", Ekologiya, Odesa, 2023.

УДК 681.391

РОЗШИРЕННЯ НЕЙРОМЕРЕЖЕВИХ МОДЕЛЕЙ АДАПТИВНИМИ ФУНКЦІЯМИ АКТИВАЦІЇ

Костюк С.
Харківський національний університет радіоелектроніки, Україна
E-mail: serhii.kostiuk@nure.ua

Анотація. *Моделі штучних нейронних мереж продовжують розвиватися у відповідь на нові завдання та виклики обробки даних. Щоб задовольнити нові вимоги до обробки даних, дослідники, як правило, розробляють нові мережеві архітектури, оновлюють існуючі та шукають можливості для покращення можливостей апроксимації для існуючих моделей. Щоб пришвидшити інновації, галузь схиляється до повторного використання попередньо навчених вагових коефіцієнтів за допомогою методів перенесення навчання. У цьому документі представлено метод розширення попередньо навчених моделей штучних нейронних мереж адаптивними функціями активації. Метод дозволяє використовувати адаптивні функції замість неадаптивних без необхідності повторного навчання. Практичні аспекти методу та його ефективність оцінюються в задачі класифікації зображень за допомогою моделі згорткової нейронної мережі на наборі даних кольорових зображень з відкритим кодом.*

Ключові слова: *адаптивна функція активації, модель нейронної мережі, трансферне навчання, глибока нейронна мережа*

Section 3. Data Security and Cryptography

УДК 004.056.5

СТЕГАНОАНАЛІТИЧНИЙ МЕТОД ВИЯВЛЕННЯ LSB-ВКЛАДЕНЬ В ЦИФРОВОМУ ВІДЕО, ПОСЛІДОВНОСТІ ЦИФРОВИХ ЗОБРАЖЕНЬ

Dr.Sci. Бобок І., Dr.Sci. Кобозєва А.

ORCID 0000-0003-4548-0709, 0000-0001-7888-0499

Національний університет «Одеська політехніка», Україна

E-mail: onu_metal@ukr.net; alla_kobozeva@ukr.net

***Анотація**. Стеганографія сьогодні є одним з напрямків захисту інформації, що найбільш швидко і ефективно розвиваються. Це призводить до підвищення актуальності стеганоаналізу, щоб уникнути негативних наслідків прихованої комунікації з метою протиправних дій. Перевага при організації прихованого каналу зв'язку сьогодні віддається цифровому відео, проте сучасний стеганоаналіз відео відстає від необхідного рівня, зокрема в умовах малої пропускної спроможності прихованого каналу зв'язку. У роботі розроблено новий ефективний стеганоаналітичний метод для виявлення результатів вбудови додаткової інформації в цифрове відео/послідовність цифрових зображень одним з найпоширеніших на сьогодні стеганографічних методів - LSB-методом. Основою розробленого методу є чутливість до збурних дій лінійності частості сингулярних векторів відповідних матриць оригінальних контентів. У роботі встановлені якісні та кількісні відмінності характеристик функції залежності частості сингулярного вектора матриці зображення/кадра відео від його номера для оригінальних контентів і стеганоповідомлень, які дали можливість для розробки алгоритмічної реалізації запропонованого методу, ефективність якої у разі малої пропускної спроможності прихованого каналу зв'язку (<0.25 біт/піксель), у тому числі при незначній кількості кадрів відео, задіяних в процесі стеганоперетворення (<50%), перевищує ефективність аналогів. Це дало можливість підвищити ефективність стеганоаналізу цифрового відео (послідовності цифрових зображень) в цілому. Показано, що запропонований метод після певної адаптації може використовуватися як загальний експертний метод для виявлення порушень цілісності цифрового відео/послідовності цифрових зображень в умовах збурних дій, зокрема незначних, що відрізняються від стеганоперетворення.*

***Ключові слова**. Цифрове відео, цифрове зображення, LSB-метод, стеганоаналіз, пропускна спроможність прихованого каналу зв'язку*

1. Вступ

Одним з основних напрямів захисту інформації, що найбільш швидко та ефективно розвивається в останні десятиліття, є стеганографія. На відміну від криптографічних засобів захисту інформації, де не існує питання про наявність приховуваного повідомлення, оскільки це є очевидним, стеганографія дозволяє скрити сам факт існування секретної інформації, вбудовуючи її в контент-контейнер, що не привертає уваги. При цьому повідомлення, що вбудовується, додатково шифрується. Таким чином, стеганографія являє собою більш високий рівень захисту інформації, в порівнянні з криптографічними засобами, що веде до того, що інтерес до стеганографії найближчим часом буде лише зростати, чому додатково сприяє стрімкий розвиток Інтернету, невід'ємною складовою якого є невирішені проблеми захисту авторських прав, прав на особисту таємницю, організації електронної торгівлі, захисту від протиправної діяльності хакерів, терористів [1]. Сьогодні стеганографія використовується для вирішення ряду різноманітних задач, серед яких [1]: захист конфіденційної інформації від несанкціонованого доступу; захист авторських прав на деякі види інтелектуальної власності; подолання систем моніторингу та керування мережевими ресурсами; створення прихованих каналів зв'язку. Взагалі комп'ютерну стеганографію сьогодні можна назвати мистецтвом прихованої комунікації. Ключовим тут є питання, у чиїх інтересах здійснюється ця комунікація. Якщо її метою є протиправні дії, результатом яких можуть бути негативні наслідки для окремих людей, підприємств, банків тощо, держави в цілому, то критично актуальним тут стає забезпечення можливості виявлення такого каналу зв'язку та отримання прихованої інформації, що знаходиться в компетенції стеганоаналізу [2]. На даний момент перевага при організації прихованого каналу зв'язку все частіше надається цифровому відео (ЦВ). Для цього існує кілька причин, основні з яких наступні: відео-трафік сьогодні становить основну частину всього споживчого інтернет-трафіку; відео дає можливість забезпечити високу пропускну спроможність прихованого каналу зв'язку, зокрема в режимі реального часу [3]; стеганоаналіз відео відстає від необхідного рівня через складність задачі, що розглядається, активність розробок тут залишається недостатньою, дослідження роблять лише початкові кроки порівняно зі стеганоаналізом окремих цифрових зображень (ЦЗ) [3]. І хоча методи, що працюють з ЦЗ, іноді адаптуються для ЦВ [4,5], така можливість є не завжди, особливо у випадку, коли стеганоаналіз ЦВ бажаний в режимі реального часу [6]. Усі стеганоаналітичні методи, у тому числі ті, що працюють з ЦВ, залежно від того, чи потрібна їм додаткова інформація про аналізований контент, можуть бути поділені на два великі класи: універсальні [7,8], що визначають наявність/відсутність додаткової інформації (ДІ) в інформаційному контенті без урахування специфіки того стеганографічного алгоритму, яким проводилася її вбудова в контейнер, і спрямовані [9], робота яких побудована з урахуванням особливостей конкретного/конкретних стеганоалгоритмів, для виявлення результатів яких вони і використовуються. Необхідно відзначити, що через широкий спектр різноманітних стеганографічних методів і підходів, що розширюється з кожним днем завдяки бурхливому розвитку стеганографії, зростає актуальність розробки універсальних стеганоаналітичних методів. Однак за наявності у стеганоаналітика необхідної інформації про організацію прихованого каналу зв'язку (зокрема, про використаний стеганоалгоритм), що на практиці має значну ймовірність, перевагу отримає відповідний спрямований метод, що має,як правило, більш високу ефективність, ніж універсальний, при застосуванні в тих же умовах. Цей висновок став основною мотиваційною складовою для авторів даної роботи для розробки саме спрямованого стеганоаналітичного методу для ЦВ.

2.Аналіз стану проблеми

Зауважимо, що формально ЦВ може розглядатися як послідовність ЦЗ-кадрів, тому в роботі такі інформаційні контенти не розрізняються, як не розрізняються ЦЗ та кадр ЦВ. Як контейнер далі розглядається ЦВ або довільна послідовність ЦЗ, збережених в однаковому форматі. Процес стеганоаналізу може проводитися в просторовій області контейнера або в одній з областей його перетворення, серед яких велика увага приділяється частотній області, області сингулярного розкладання відповідної матриці тощо. Важливим і бажаним у процесі стеганоаналізу є те, щоб ефективність експертизи контенту в обраній для неї області контейнера не залежала від того, яка саме область використовувалася при вбудові ДІ [3,7,10]. Одним із найпоширеніших стеганографічних методів на сьогоднішній день залишається метод модифікації найменшого значущого біта (*Least Significant Bit*) – LSB-метод [11]. Простота його реалізації, потенційна можливість забезпечення високої пропускної спроможності прихованого каналу зв'язку (ПСПК), можливість здійснення вкладення та декодування ДІ в режимі реального часу, очевидно, дасть можливість цьому методу не втратити свою затребуваність у найближчі роки, залишаючи актуальною задачу ефективного виявлення вкладень ДІ, проведених цим методом. Питання виявлення результатів стеганоперетворення ЦВ LSB-методом вже неодноразово розглядалися фахівцями. Так в [12] запропоновано метрику, яка використовується в процесі стеганоаналізу ЦВ, що дозволяє виявляти збурення в кадрі ЦВ в областях з малими перепадами значень яскравості (фонових областях). Основна перевага цього методу досягається завдяки врахуванню напряму та відстані поточного кадру ЦВ від опорного. В [13] запропоновано покадровий метод стеганоаналізу ЦВ, розрахований на виявлення вкладень ДІ, зроблених LSB-методом (реалізація LSB-*matching*), який став «відповіддю» авторів на стандарт кодування відео H265. Наведені авторами в роботі докладні експериментальні результати, у тому числі при використанні малої ПСПК, показують ефективність запропонованого методу відеостеганоаналізу, та разом з умовами застосування (покадровий аналіз, LSB-matching-реалізація, а головне – дієздатність при малій ПСПК, що у більшості робіт навіть не досліджується) робить його пріоритетним для порівняльного аналізу ефективності з методом, запропонованим у роботі. В [14] запропоновано методику стегоаналізу ЦВ на основі крос-кореляційного аналізу, де увага приділяється виявленню результатів вбудови ДІ за допомогою реалізації LSB-matching. В основі [5] лежить аналіз кількості блоків, в яких пікселі мають однакові значення яскравості в межах блоку в матрицях кольорових складових цифрових контентів. Основною перевагою тут є використання в процесі стеганоаналізу просторової області ЦВ, що дозволяє уникнути додаткових обчислювальних витрат та додаткового накопичення обчислювальної похибки при переході з просторової області в область перетворення і назад. Однак жоден із згаданих вище стеганоаналітичних методів не дозволяє говорити про задовільне остаточне вирішення задачі стеганоаналізу ЦВ, зокрема в умовах малої ПСПК. Нещодавно авторами роботи було розроблено новий підхід до проблеми виявлення порушення цілісності ЦЗ, заснований на аналізі функції залежності частості сингулярного вектора (СНВ) матриці зображення від номера (СНВ-підхід) [15]. Основою СНВ-підходу є встановлена властивість лінійності частості СНВ для оригінального ЦЗ. Ця властивість є чутливою до збурних дій, а тому дає принципову можливість для відокремлення оригінального контенту від такого, цілісність якого порушена, в тому числі в умовах малих збурних дій. З урахуванням можливості розгляду ЦВ як сукупності ЦЗ, а процесу стеганоперетворення як окремого випадку збурної дії [7], стає можливим і перспективним використання СНВ-підходу для розробки стеганоаналітичних методів для аналізу ЦВ. СНВ-підхід з урахуванням згаданої чутливості лінійності частості СНВ має забезпечити ефективність відповідних методів, зокрема в умовах малої ПСПК.

3. Мета та задачі дослідження

Метою роботи є підвищення ефективності стеганоаналізу ЦВ/послідовності ЦЗ шляхом розробки стеганоаналітичного методу для виявлення результатів вбудови ДІ LSB-методом, ефективного, у тому числі, при малій ПСПК, на основі СНВ-підходу. Під малою пропускною спроможністю прихованого каналу зв'язку далі розуміється ПСПК<0.25 біт/піксель, обгрунтування чого можна знайти в [4]. Для досягнення поставленої мети в роботі розв'язуються наступні задачі: вибір кількісного показника порушення лінійності частості СНВ; дослідження кількісних характеристик функції залежності частості СНВ матриці зображення/кадра ЦВ від його номера для оригінальних контентів та стеганоповідомлень, отриманих за допомогою LSB-методу з використанням різної ПСПК; розробка стеганоаналітичного методу для ЦВ та його алгоритмічної реалізації; аналіз ефективності розробленого методу.

4. Розробка стеганоаналітичного методу

4.1. Поняття частості сингулярного вектора

Розглянемо основи СНВ-підходу [15]. Нехай F - $n\times n$-матриця ЦЗ/кадра ЦВ,

$$F = U\Sigma V^T = \sum_{i=1}^{n} \sigma_i u_i v_i^T \qquad (1)$$

- нормальне сингулярне розкладання F, що визначається однозначно [16], де U, V – ортогональні $n \times n$-матриці, стовпці яких u_i, v_i, $i = \overline{1,n}$, - ліві й праві СНВ F відповідно, при цьому ліві СНВ додатково є лексикографічно додатними [16]; $\Sigma = diag(\sigma_1,..., \sigma_n)$, $\sigma_1 \geq ... \geq \sigma_n \geq 0$ – сингулярні числа (СНЧ) F. Використання лише одної матриці для формального представлення ЦЗ не обмежує спільність міркувань. У разі кольорового ЦЗ як F може виступати матриця будь-якої (чи послідовно всіх) колірної складової (схема RGB), матриця яскравості (схема YUV). СНВ u_i (v_i), елементи якого в межах СНВ-підходу розглядаються як значення деякої дискретної функції на проміжку [1,n], ставиться у відповідність кількісна характеристика – частість $\bar{\eta}$, що визначається наступним чином [15]:

$$\bar{\eta} = \begin{cases} \dfrac{\eta}{2}, \text{ якщо } \eta - \text{парне} \\ \dfrac{\eta+1}{2}, \text{ якщо } \eta - \text{непарне} \end{cases} \qquad (2)$$

де η - число змін знака функції на розглянутому проміжку.

Позначимо $fr(i), i = \overline{1,n}$, - функцію, що відображає залежність частості СНВ u_i (v_i) від його номеру i.

В результаті дослідження властивостей $fr(i)$ встановлено [15], що для оригінальних ЦЗ (у форматі без втрат) практично постійною є швидкість зростання тренду цієї функції практично на всьому сегменті [1, n], наслідком чого є властивість, яка отримала назву *лінійності частості СНВ* (як лівих, так і правих) для ЦЗ/кадра ЦВ, цілісність якого не порушена, що означає, що функція $fr(i), i = \overline{1,n}$, добре апроксимується практично на всьому сегменті [1, n] лінійною монотонно зростаючою функцією $l(i), i = \overline{1,n}$ для оригінального ЦЗ без втрат (рис.1(а,б)). Лінійність частості СНВ є характеристикою, чутливою до будь-яких збурних дій, зокрема до стеганоперетворення (рис.1(в)).

Порушення лінійності частості СНВ для неоригінального ЦЗ/кадра ЦВ, що зазнало вбудову ДІ, дозволить відокремлювати його від зображення/кадра ЦВ, цілісність якого порушена не була (частини незаповненого контейнера), при цьому відповідно до СНВ-підходу така принципова можливість може бути ефективно реалізована і у разі малої ПСПК.

Однак для розробки відповідного стеганоаналітичного методу необхідно отримати кількісні показники порушення лінійності частості СНВ, що є одною з задач роботи.

Зауважимо, що, враховуючи нестійкість LSB-методу до будь-яких атак проти вбудованого повідомлення, зокрема до атаки стиском, одержувані стеганоповідомлення будуть зберігатися у форматах без втрат, при цьому контейнери можуть бути як у форматі з втратами, так і у форматі без втрат, що потребує окремого дослідження.

4.2. Кількісні показники порушення лінійності частості сингулярного вектора

Як можливі кількісні показники порушення лінійності частості СНВ для кадру ЦВ/ЦЗ розглянемо наступні:

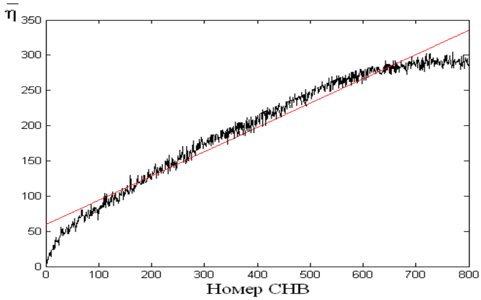
$$P = \left(\sum_{i=1}^{n} \left(fr(i) - l(i) \right)^2 \right)^{\frac{1}{2}}, \quad (3)$$

$$L = \max_i \left| fr(i) - l(i) \right| \cdot \quad (4)$$

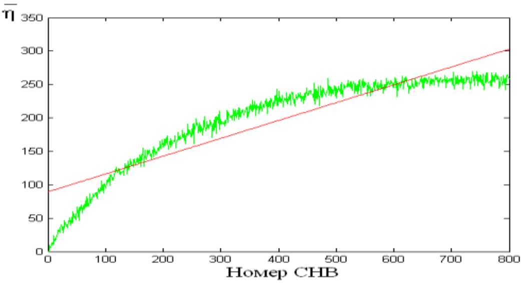
Розглянемо показники (3) та (4) для окремих ЦЗ (кадрів ЦВ) залежно від того, оригінальним або неоригінальним є контент (частина контейнера/частина стеганоповідомлення), від формату контенту (із втратами, без втрат), від величини ПСПК.

Виходячи з вищенаведеного, теоретично природним показником порушення цілісності контенту буде збільшення як (3), так і (4), порівняно з оригінальним ЦЗ/кадром оригінального відео, що зберігаються у форматі без втрат (збереження з втратами розглядається як збурна дія), за будь-яких змін, у тому числі в результаті стеганоперетворення.

а



б

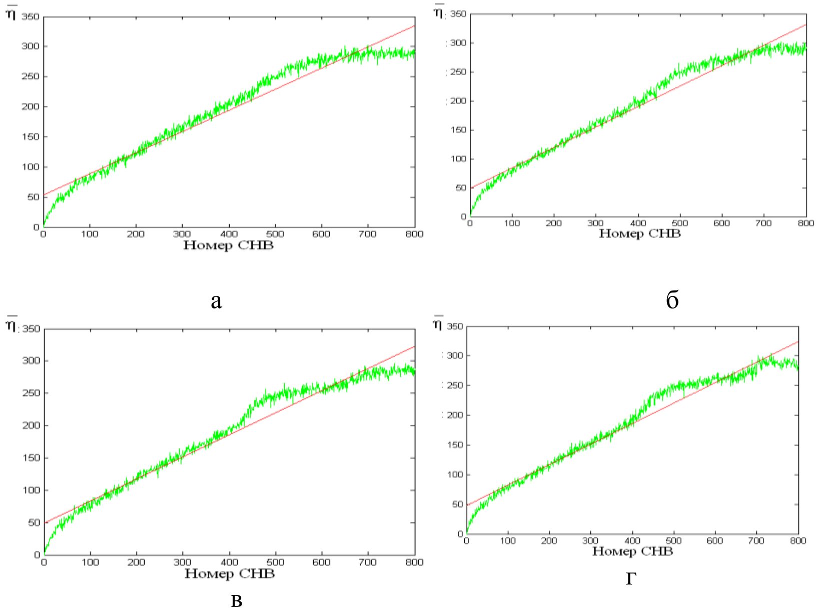


в

Рис.1. Ілюстрація лінійності частості СНВ для оригінального ЦЗ та її порушення у разі стеганоповідомлення: а – оригінальне ЦЗ; б – графіки функції $fr(i)$ і її лінійної проксимації $l(i)$ для оригінального ЦЗ; в – графіки функції $fr(i)$ і $l(i)$ для ЦЗ-стеганоповідомлення, сформованого LSB-методом з ПСПК=0.5 біт/піксель

Але на практиці тут можливі деякі «винятки»: показники (3) і (4) хоч і введені природним чином, однак, на жаль, як показує обчислювальний експеримент, не забезпечують повною мірою можливість відокремити оригінальний контент (окреме ЦЗ, окремий кадр ЦВ) від такого, цілісність якого порушена, з урахуванням лише їх числових значень, хоча таке відокремлення не викликає питань при експертизі безпосередньо графіків функцій $fr(i)$ і $l(i)$, де очевидним є порушення лінійності частості СНВ, що проілюстровано (рис.2) для ЦЗ, представленого на рис.1(а), для якого $P = 531, L = 35$, при тому, як для збурених стисненням ЦЗ ці показники можуть бути меншими (рис.2). Однак, як показує обчислювальний експеримент, кількість таких контентів (окремих кадрів ЦВ, окремих ЦЗ в послідовності), для яких збільшення показників (3), (4) при порушеннях цілісності не відбувається, невелика – менше 7% при різних збурних діях, а у разі стеганоперетворення LSB-методом вони практично відсутні. З огляду на це параметри (3), (4) використовуються далі для кількісної оцінки порушення лінійності частості СНВ.

Розглянемо значення параметрів (3), (4) для оригінальних ЦЗ в форматі без втрат (для визначеності – Tif). З огляду на характерні особливості функції $fr(i)$ (відсутність монотонності на області визначення) очевидно, що для будь-якого ЦЗ/кадра ЦВ, у тому числі оригінального, значення $P \neq 0, L \neq 0$.

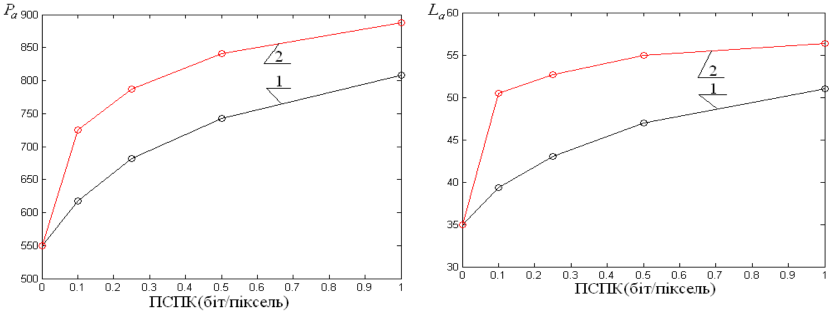


а б в г

Рис.2. Графіки функцій $fr(i)$ *і* $l(i)$ *для ЦЗ, що зазнало стиск з втратами: а – QF=85 (* $P = 496, L = 33$ *); б – QF=75 (* $P = 441, L = 41$ *); в – QF=65 (* $P = 440, L = 43$ *); г – QF=55 (* $P = 431, L = 37$ *)*

В результаті обчислювального експерименту, в якому було задіяно понад 600 ЦЗ з бази img_Nikon_D70s (*n*=800) [17] встановлено, що середнє значення показника *P*, що далі позначається P_a, визначається як $P_a = 549.82$, а аналогічне значення L_a для показника *L* дорівнює 34.97. На основі цих зображень були сформовані групи стеганоповідомлень методом LSB-matching з ПСПК=0.1, 0.25, 0.5, 1 біт/піксель. Для кожної з отриманих групп визначалися P_a, L_a (рис.3). Графіки залежності P_a, L_a від величини ПСПК поводяться відповідно до теоретичних очікувань, монотонно зростаючи: зі збільшенням ПСПК – сили збурної дії на оригінальне ЦЗ збільшується реакція $fr(i)$, що виражається в збільшенні відмінності значень $fr(i)$ від значень відповідної лінійної апроксимації. Оскільки стиск ЦЗ з втратами є додатковою до стеганоперетворення збурною дією на оригінальний контейнер без втрат, то при використанні в якості контейнера контенту у форматі з втратами очевидним є збільшення відмінності P_a, L_a для стеганоповідомлень від значень цих же параметрів для оригінальних ЦЗ, що знайшло своє відображення на рис.3. Виходячи з отриманих експериментальних результатів, значення P_a, L_a значно відрізняються для послідовностей оригінальних ЦЗ від послідовностей ЦЗ, що представляють з себе стеганоповідомлення, сформовані LSB-matching з наступним збереженням в форматі без втрат, про необхідність чого було згадано вище. Ця відмінність має місце незалежно від того, в якому форматі фігурує контейнер – з втратами чи без втрат. Тут дуже важливою є явна кількісна відмінність P_a, L_a при малій ПСПК=0.1 біт/піксель, що є теоретично очікуваним наслідком чутливості лінійності частості СНВ оригінального контенту без втрат до збурних дій. Враховуючи, що відмінності параметрів P_a і L_a для оригінальних і неоригінальних контентів порівнянні між собою (при ПСПК=0.1 біт/піксель для P_a збільшення при вбудові ДІ склало 11.5%, для L_a -

12.6%), для подальших досліджень прийнято рішення залишити лише один із двох розглянутих параметрів, а саме (4).



а б

Рис.3. Графіки залежності P_a, L_a от ПСПК: 1 – для контейнерів в форматі без втрат; 2 – для контейнерів в форматі з втратами (Jpeg, QF=85)

Середнє значення L_a несе в собі інформацію про значення всієї сукупності, для якої воно визначається, проте для мети, поставленої в роботі, з урахуванням того, що при стеганоперетворенні ЦВ/послідовності ЦЗ для вбудови ДІ може використовуватися не кожен кадр ЦВ (не кожне ЦЗ послідовності) [13], важливою є також інформація про те, наскільки часто в послідовності ЦЗ/кадрів ЦВ приймаються однакові/близькі значення *L*, яке з можливих значень приймається найчастіше. Для отримання такої інформації скористаємося гістограмами значень *L*, отриманих для окремих кадрів ЦВ чи окремих ЦЗ послідовності. Виходячи з теоретичних передумов (СНВ-підхід) та результатів, що стосуються L_a, можна припустити, що гістограми значень *L* для контейнера - оригінальної (в форматі без втрат) послідовності кадрів ЦВ або ЦЗ та для стеганоповідомлення - неоригінальної послідовності, що є результатом вбудови ДІ в контейнер, якісно відрізнятимуться одна від іншої:

– гістограма Γ_S для послідовності кадрів/ЦЗ стеганоповідомлення при будь-якій ПСПК буде зсунута вправо вздовж осі *OL* відносно гістограми Γ_O для оригінальної послідовності (тим більше, чим більше ПСПК);

– моди $m(\Gamma_S), m(\Gamma_O)$ гістограм Γ_S і Γ_O відповідно будуть пов'язані співвідношенням:

$$m(\Gamma_S) \geq m(\Gamma_O); \quad (5)$$

– із зростанням ПСПК відмінність (5) між $m(\Gamma_S), m(\Gamma_O)$ буде збільшуватися через зростання $m(\Gamma_S)$.

Всі теоретичні очікування знайшли своє підтвердження на практиці (рис.4). Для розробки стеганоаналітичного методу необхідно встановити кількісні відмінності для $m(\Gamma_S), m(\Gamma_O)$.

Найчастіше на практиці при використанні в якості контейнера ЦВ або послідовності ЦЗ вбудова ДІ відбувається в окремі кадри/ЦЗ з незмінною ПСПК.

Тоді, отримуючи відповідні гістограми, можна не тільки відокремити оригінальне ЦВ від такого, яке несе в собі приховану ДІ, а й оцінити ПСПК, використовуючи безпосереднє значення моди (рис.4), що є напрямом подальшої роботи авторів.

Але якщо ПСПК і змінювалася під час вбудовування ДІ від кадру до кадру, це, враховуючи основи СНВ-підходу, не завадить виявити факт наявності ДІ за значенням моди відповідної гістограми, що ілюструють результати, представлені на рис.5(а), де мода дорівнює 43, тоді як для оригінальної послідовності кадрів/ЦЗ це значення набагато менше (рис.4(а)).

Найбільш часто на практиці як контейнер використовується ЦВ/послідовність ЦЗ у форматі з втратами, які після вбудови ДІ зберігаються у форматі без втрат, враховуючи нестійкість методу LSB до будь-яких атак проти вбудованого повідомлення.

У цьому випадку відмінність у модах гістограм значень L для первинної оригінальної послідовності (у форматі без втрат) і стеганоповідомлення, побудованого на основі контейнера, отриманого перезбереженням первинної послідовності у формат із втратами, буде ще очевиднішим.

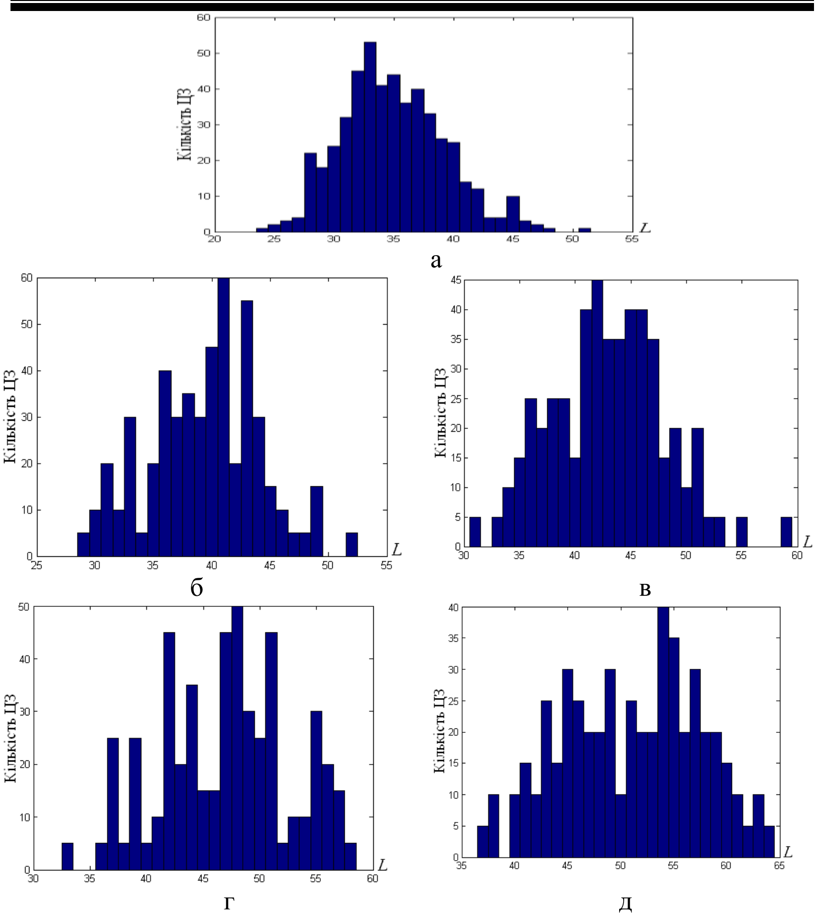
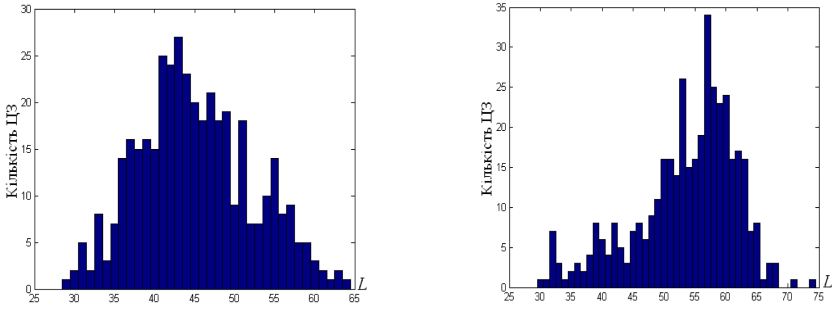


Рис.4. Гістограми значень L: а – для контейнера, що є послідовністю 500 оригінальних ЦЗ в форматі Tif; б, в, г, д - для стеганоповідомлень, сформованих методом LSB з ПСПК=0.1, 0.25, 0.5, 1 біт/піксель відповідно

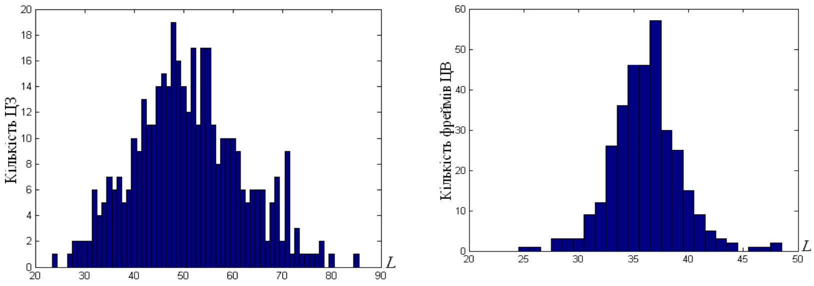
Справді, саме отримання контейнера у форматі із втратами вже порушить лінійність частості СНВ ЦЗ/кадра ЦВ первинної послідовності (рис.6 (порівн. з рис.4(а))), а стеганоперетворення лише посилить цей відступ від лінійності (рис.7).



а б

Рис.5. Гістограма значень L окремих ЦЗ для стеганоповідомлення, сформованого методом LSB (ПСПК приймала різні значення для різних ЦЗ: 0.1, 0.25, 0.5, 1 біт/піксель), отриманого на основі контейнера, що є послідовністю з 400 ЦЗ в форматі: а – Tif; б – Jpeg (QF=85)

При зміні ПСПК від кадру до кадру значна перевага $m(\Gamma_S)$ в порівнянні з $m(\Gamma_O)$ не викликає проблем у виявленні стеганоповідомлення (рис.5(б)).



а б

Рис.6. Гістограма значень L: а - для послідовності відповідних ЦЗ в форматі Jpeg ($QF \in \{55,65,75,85\}$); б – для ЦВ в форматі Mpeg4

Необхідно відзначити, що у випадку, коли контейнер являє собою ЦВ/послідовність ЦЗ у форматі з втратами, неможливо систематично оцінювати за значенням $m(\Gamma_S)$ величину ПСПК навіть у тому випадку, коли вона не змінюється від кадру до кадру.

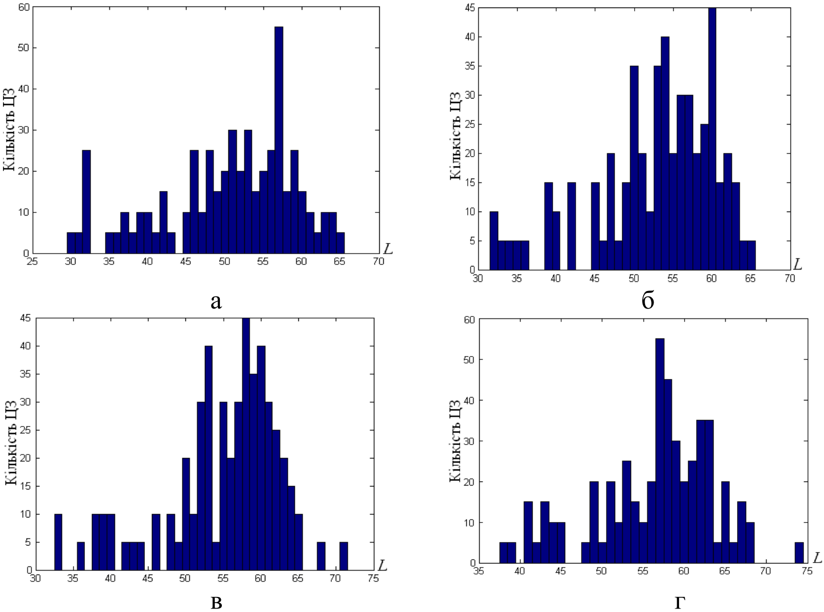


Рис.7. Гістограми значень L окремих ЦЗ для стеганоповідомлення, сформованого на основі контейнера, що є послідовністю з 500 ЦЗ в форматі Jpeg (QF=85) методом LSB з ПСПК: а – 0.1 біт/піксель; б – 0.25 біт/піксель; в – 0.5 біт/піксель; г – 1 біт/піксель

Це є наслідком «накладання» двох збурних дій на оригінальний контент у форматі без втрат: стиснення і стеганоперетворення LSB-методом. Їхній вплив на сингулярні трійки (σ_i, u_i, v_i) (1) матриці F відбувається незалежно один від одного, десь посилюючи, а десь зменшуючи впливи один одного, результатом чого і може бути немонотонність зростання $m(\Gamma_S)$, що спостерігається на рис.7, хоча для переважної більшості розглянутих ЦВ у форматі із втратами монотонне зростання $m(\Gamma_S)$ зі зростанням ПСПК все ж таки мало місце (результати експерименту наведені нижче). Зауважимо, що незважаючи на те, що більшість отриманих вище результатів продемонстровано на послідовностях ЦЗ, в силу розгляду авторами ЦВ як сукупності окремих кадрів-ЦЗ, а аналізу ЦВ як послідовного аналізу кадрів, це ніяк не обмежує спільності міркувань та висновків, підтвердженням чого є результати проведених експериментів, деякі з яких представлені в табл.1.

Таблиця 1

Кількісні показники значень *L*, $m(\Gamma_S)$ при різних значеннях ПСПК (вбудова ДІ в кожний кадр ЦВ)

ЦВ/послідовність ЦЗ	ПСПК (біт/піксель)	*L*	$m(\Gamma_S)$
V1 (100 кадрів)	0.1	42.12	40
	0.25	47.10	46
	0.5	51.33	51
	1	54.47	52
V2 (100 кадрів)	0.1	41.61	41
	0.25	45.84	47
	0.5	50.50	50
	1	54.36	54
V3 (100 кадрів)	0.1	40.42	41
	0.25	44.72	47
	0.5	50.27	51
	1	56.90	54
V4 (325 кадрів)	0.1	41.35	41
	0.25	46.98	46
	0.5	50.68	50
	1	55.60	54
V5 (1000 кадрів)	0.1	57.24	56
	0.25	62.74	61
	0.5	64.66	63
	1	66.04	65
I1 (100 ЦЗ)	0.1	50.46	57
	0.25	52.64	60
	0.5	54.94	58
	1	56.36	57
I2 (100 ЦЗ)	0.1	39.37	41
	0.25	43.05	42
	0.5	46.95	48
	1	51.01	54

V1-V4 – ЦВ (MPEG-4), отримані мобільним телефоном Realme 7 pro (64 MP, f/1.8, 26mm (wide), 1/1.73", 0.8µm, PDAF), V5 – ЦВ (MPEG-4), отримане відеокамерою Canon PowerShot A520 – CCD, 4MP; I1 (формат Jpeg (*QF*=85)), I2 (формат Tif) – послідовності ЦЗ, де I2 містить зображення з бази [17], а I1 отримана перезбереженням ЦЗ з I2 в формат з втратами.

4.3. Стеганоаналітичний метод та його алгоритмічна реалізація

Враховуючи, що властивості відповідних лівих і правих СНВ матриці F ЦЗ/кадра ЦВ в межах СНВ-підходу якісно не відрізняються, а кількісно є порівнянними, у запропонованому методі аналізуються тільки ліві лексикографічно додатні СНВ.На основі отриманих результатів основні кроки стеганоаналітичного методу для експертизи ЦВ, що містить N кадрів, або послідовності N ЦЗ, формально збережених у форматі без втрат, виглядають наступним чином.

Крок 1. Для кожного з N кадрів ЦВ/кожного ЦЗ послідовності з матрицею $F_j,\ j=\overline{1,N}$, розміром $n\times n$:

1.1. Побудувати нормальне сингулярне розкладання (1).

1.2. Визначити частості (2) кожного лівого СНВ $u_i, i=\overline{1,n}$, матриці F_j кадра ЦВ/ЦЗ послідовності.

1.3. Побудувати функцію $fr(i),\ i=\overline{1,n}$, залежності частості СНВ u_i від його номеру *i*.

1.4. Побудувати лінійну апроксимацію $l(i)$ для функції $fr(i)$.

1.5. Визначити значення *L* для F_j - $L(F_j)$ (4).

Крок 2. Визначити

$$L_a=\frac{1}{N}\sum_{j=1}^{N}L(F_j).$$

Крок 3. Побудувати гістограму Γ значень $L(F_j),\ j=\overline{1,N}$.

Крок 4. Визначити моду $m(\Gamma)$ отриманої на попередньому кроці гістограми Γ.

Крок 5. *Якщо*

$$(m(\Gamma)<T_1)\&(L_a<T_2),$$

де T_1, T_2 - порогові значення, отримані експериментально,
то ЦВ/послідовність ЦЗ не містить ДІ
інакше ЦВ/ послідовність ЦЗ є стеганоповідомленням.

Алгоритмічній реалізації методу відповідають наступні значення параметрів: *n*=800, $T_1 = 35$, $T_2 = 36.5$.

Ефективність розробленого методу оцінювалася за допомогою стандартного для таких задач параметра – точності виявлення порушення цілісності [18] (*accuracy* (*ACC*)):

$$ACC = (TP + TN)/(TP + FN + TN + FP), \quad (6)$$

де *TP* (*True Positive*) — кількість правильно виявлених стеганоповідомлень; *TN* (*True Negative*) — кількість правильно виявлених контейнерів; *FP* (*False Positive*) — кількість контейнерів, що помилково визначені як стеганоповідомлення (помилки II роду); *FN* (*False Negative*) - кількість пропущених стеганоповідомлень (помилки I роду). Для оцінки ефективності алгоритмічної реалізації розробленого стеганоаналітичного методу було проведено обчислювальний експеримент. В експерименті було задіяно 35 контейнерів – ЦВ та послідовностей ЦЗ, кількість кадрів/ЦЗ у яких була від 100 до 1000 одиниць. ЦВ були отримані кількома відеокамерами (Olympus SP-820 – CMOS, 14MP; Nikon COOLPIX P100 – CMOS, 10MP; Canon PowerShot A520 – CCD, 4MP; відеокамерой мобільного телефону Realme 7 pro (64 MP, f/1.8, 26mm (wide), 1/1.73", 0.8µm, PDAF)). Послідовності ЦЗ в форматі без втрат компанувалися зображеннями з бази img_Nikon_D70s [18], перезбереженням яких в формат Jpeg з $QF \in \{55, 65, 75, 85\}$ отримувалися послідовності ЦЗ в форматі з втратами. Для кожного з описаних 35 контентів були сформовані стеганоповідомлення методом LSB-matching, де ПСПК для кадру ЦВ/ЦЗ послідовно бралася 0.1, 0.25, 0.5, 1 біт/піксель. Вбудова ДІ проводилася у 100, 80, 60, 40% кадрів ЦВ/ЦЗ послідовності. Таким чином, загальна кількість стеганоповідомлень склала 560. Результати оцінки ефективності методу з використанням показника (6), що включають порівняння з аналогом [13], наведені в табл.2,3. Як випливає з наведених результатів (табл.2), розроблений метод забезпечує високу ефективність за умов вбудови ДІ у кожен кадр ЦВ, дозволяючи підвищити цю ефективність на 1.5% при ПСПК=0.1 біт/піксель. При цьому при тестуванні розробленого методу помилки 1 роду виявлені не були, що є його значною перевагою порівняно з аналогами [5,13].

Таблиця 2

Результати порівняльного аналізу ефективності, що визначається показником *ACC* (%), розробленого методу при вбудові ДІ у всі кадри ЦВ/ЦЗ послідовності

Метод	ПСПК (біт/піксель)			
	1	0.5	0.25	0.1
Запропонований метод	**98.86**	**98.86**	**98.86**	**98.86**
Метод [13]	99.95	99.91	99.88	97.44

Таблиця 3

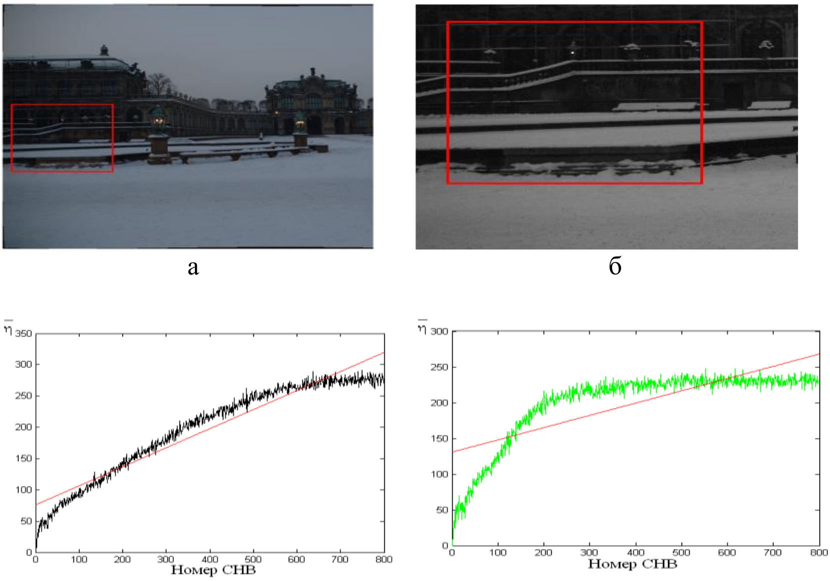
Результати порівняльного аналізу ефективності, що визначається показником *ACC* (%), розробленого методу залежно від кількості кадрів ЦВ, задіяних в процесі стеганоперетворення

ПСПК (біт/піксель)	Кількість кадрів ЦВ, задіяних в процесі стеганоперетворення (%)					
	80		60		40	
	Метод [13]	Розроб-лений метод	Метод [13]	Розроб-лений метод	Метод [13]	Розроб-лений метод
0.5	100	**98.86**	99.78	**98.86**	100	**98.86**
0.25	100	**98.86**	100	**98.86**	99.26	**96.00**
0.1	93.81	**98.86**	92.76	**98.86**	79.74	**91.43**

Чутливість властивості лінійності частості СНВ дала можливість отримати для розробленого стеганоаналітичного методу великий виграш в ефективності порівняно з аналогами при малій ПСПК: максимально показник *ACC* підвищений на 11.69% в умовах кількості кадрів ЦВ, задіяних в процесі стеганоперетворення, 40% (табл.3).

Зауваження 1. Одним з найпріоритетніших напрямків захисту інформації сьогодні є виявлення будь-яких порушень цілісності цифрових контентів. Математична основа запропонованого стеганоаналітичного методу (СНВ-підхід) робить його перспективним для виявлення результатів порушення цілісності ЦВ/послідовності ЦЗ, що відбулося при збурній дії, яка відрізняється від стеганоперетворення. Це відбувається через встановлену чутливість лінійності частості СНВ до довільних збурних дій [15]. Практичним підтвердженням зроблених висновків є результати обчислювального експерименту, де були задіяні послідовності ЦЗ, в результаті якого встановлено, що при таких збурних діях, як накладання різноманітних шумів, а саме: гауссівського шуму з нульовим математичним очікуванням і $D \in \{0.01, 0.01, 0.001, 0.0001, 0.00001\}$, мультиплікативного шуму з $D \in \{0.001, 0.0001, 0.00001\}$, пуассонівського шуму значення параметра *ACC* було близьким до 99%, при цьому важливо зазначити, що ефективність запропонованого методу не знижується в умовах навіть зовсім незначних збурних дій. Так в умовах накладання гауссівського шуму з $D = 0.00001$незначні зміни (не більше, ніж на 5 градацій яскравості) зазнають лише 11-26% пікселів ЦЗ; для мультиплікативного шуму з $D = 0.00001$ зміни яскравості пікселей (4-47% загальної кількості пікселів) відбуваються лише на 1-2 градації. Порушення цілісності, що відбувалося шляхом накладання гауссівського шуму з $D = 0.0000015$(зміна на 1-2 градації яскравості лише 3.8-5.6% загальної кількості пікселів ЦЗ) призвело до *ACC*=85.2%. Для ілюстрації спроможності розробленого метода в умовах збурних дій, що відрізняються від результатів LSB-перетворення, накладання різноманітних шумів обрано не випадково, оскільки виявлення шумів є вкрай актуальною задачею з урахуванням того, що результат будь-якої збурної дії загалом може розглядатися як накладання деякого шуму [19,20]; крім того, шум часто використовується для маскування клонування, стеганоперетворення.

Зауваження 2. Використання СНВ-підходу є перспективним і в тому випадку, коли збурні дії застосовуються лише до деякої зв'язної частини ЦЗ. Такий спосіб зміни ЦЗ є актуальним, зокрема при вбудові цифрового водяного знаку, який може розташовуватися в деякій порівняно незначній області зображення. Використання СНВ-підходу тут потребує розбивки матриці досліджуваного контенту на непересічні $m \times m$-блоки, дозволяє «грубо» локалізувати неоригінальну область ЦЗ/кадру ЦВ (рис.8). Ефективність та точність локалізації неоригінальної області тут буде залежати від того, наскільки розміри зміненої області будуть відрізнятися від розміру блоків при розбивці матриці контенту. Максимальна ефективність буде досягатися у випадку порівнянності лінійних розмірів зміненої області з m; мінімальна ефективність буде відповідати випадку, коли лінійні розміри будуть значно меншими m.



в г

Рис.8. Ілюстрація дієздатності СНВ-підходу при виявленні локальних порушень цілісності: а – ЦЗ, цілісність якого порушена локально шляхом накладання гауссівського шуму з нульовим математичним очікуванням і D=0.0001 на область, що обмежена червоним прямокутником; б – частина ЦЗ, для якої зафіксовано порушення лінійності частості СНВ; в – типовий графік функції $fr(i)$ *для підобласті ЦЗ, цілісність якої не порушена; г - графік функції* $fr(i)$ *для підобласті ЦЗ (рис.8(б)), цілісність якої порушена*

Встановлення кількісних значень цієї відмінності в межах адаптації розробленого методу для виявлення локальних порушень цілісності цифрових контентів – задача, на яку спрямовані зусилля авторів в даний час.

Зауваження 3. Локальні порушення цілісності ЦЗ можуть відбуватися шляхом застосування таких інструментів, як клонування [21] чи фотомонтаж [22]. Враховуючи, що на практиці при використанні для зміни ЦЗ згаданих програмних інструментів необхідною є додаткова обробка зображення для «адаптації» чужерідних частин в нових для них областях зображення (фільтрація, розмиття, розмиття по контуру клону, лінії з’єднання частин різних ЦЗ при фотомонтажі, корекції кольору, яскравості тощо), а також те, що при фотомонтажі часто відбувається використання зображень, що мають різні формати збереження (з/без втрат), застосування запропонованого методу тут дозволить виявляти результати порушення цілісності. Але якщо згадані локальні порушення будуть відбуватися з використанням частин лише оригінальних ЦЗ, то застосування запропонованого методу може виявитися неефективним.

5. Висновки

У роботі вирішено важливу науково-практичну задачу підвищення ефективності виявлення стеганографічного каналу зв'язку у випадку, коли в якості контейнера використовується ЦВ, послідовність ЦЗ. На базі СНВ-підходу розроблено ефективний стеганоаналітичний метод для виявлення результатів вбудови ДІ в ЦВ/послідовність ЦЗ LSB-методом, основою якого є чутливість лінійності частості СНВ до збурних дій. У ході розробки були встановлені кількісні відмінності характеристик функції залежності частості СНВ матриці зображення/кадру ЦВ від його номера для оригінальних контентів і стеганоповідомлень, що отримані за допомогою LSB-методу з використанням різної ПСПК. Як такі характеристики в результаті обґрунтованого вибору виступили: середнє значення по всіх кадрах ЦВ/усім ЦЗ послідовності максимального відхилення функції $fr(i)$, що визначає частість СНВ залежно від його номера, від її лінійної апроксимації, а також мода гістограми максимальних відхилень $fr(i)$ від лінійної апроксимації, отриманих для кожного кадру ЦВ/кожного ЦЗ послідовності. Встановлені кількісні відмінності дали можливість розробки алгоритмічної реалізації запропонованого методу, ефективність якої у разі малої ПССК перевищує ефективність аналога, що визначається за допомогою параметра *ACC*, при цьому максимальне збільшення відповідає ПССК=0.1 біт/піксель за умов кількості кадрів ЦВ, задіяних в процесі стеганоперетворення, 40% і складає більше 11%, що є основним практичним значенням результатів роботи і є основним показником підвищення ефективності стеганоаналізу ЦВ, послідовності ЦЗ в цілому. В даний час зусилля авторів спрямовані на адаптацію розробленого методу до роботи в умовах малої кількості кадрів ЦВ, ЦЗ в послідовності (10 і менше), де знижується інформативність гістограм, що розглядаються. Математичні основи запропонованого методу дають принципову можливість використовувати його не тільки як стеганоаналітичний для виявлення LSB-вкладень, але і як загальний експертний метод після певної адаптації для виявлення порушень цілісності ЦВ/послідовності ЦЗ в умовах збурних дій, зокрема незначних, що відрізняються від стеганоперетворення, що є актуальною сучасною проблемою в галузі захисту інформації.

Література

1. В.О.Хорошко, Ю.Є.Яремчук, В.В.Карпінець, Комп'ютерна стеганографія. Вінниця: ВНТУ, 2017.

2. W.M. Eid, S.S. Alotaibi, H.M. Alqahtani, S.Q. Saleh, Digital Image Steganalysis: Current Methodologies and Future Challenges, IEEE Access 10 (2022) 92321-92336. DOI: 10.1109/ACCESS.2022.3202905.

3. M. Bouzegza, A. Belatreche, A. Bouridane, M. Tounsi, A comprehensive review of video steganalysis, IET Image Process. 16 (2022) 3407–3425. DOI:10.1049/ipr2.12573.

4. I.I. Bobok, Steganalysis method for detection of the hidden communication channel with low capacity, Telecommunications and Radio Engineering 77 (2018) 1597–1604. DOI:10.1615/TelecomRadEng.v77.i18.20.

5. A.V. Akhmametieva, A.A. Efimenko, Comparative efficiency analysis of steganalytic algorithms for detecting the presence of attachments of confidential information in digital video, Informatics and Mathematical Methods in Simulation 7 (2017) 96-102.

6. K. Tasdemir, F. Kurugollu, S. Sezer, Spatio-Temporal Rich Model-Based Video Steganalysis on Cross Sections of Motion Vector Planes, IEEE Transactions on Image Processing 25 (2016) 3316-3328. DOI: 10.1109/TIP.2016.2567073.

7. A.A. Kobozeva, I.I. Bobok, A.I. Garbuz, General Principles of Integrity Checking of Digital Images and Application for Steganalysis, Transport and Telecommunication Journal 17 (2016) 128-137. DOI:10.1515/ttj-2016-0012.

8. M. Dalal, M. Juneja, Steganography and Steganalysis (in digital forensics): a Cybersecurity guide. Multimed Tools Appl 80 (2021) 5723–5771. DOI:10.1007/s11042-020-09929-9.

9. R. Rasras, M. Abu Sara, Z. AlQadi, R. Abu zneit, Comparative Analysis of LSB, LSB2, PVD Methods of Data Steganography, International Journal of Advanced Trends in Computer Science and Engineering 8 (2019) 748-754.
10. S. Li, X. Ma, P. Liu, Q. Dai, H. Deng, Detection of Information Hiding by Modulating Intra Prediction Modes in H.264/AVC, in: Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering (ICCSEE 2013), 2013, pp. 590-593. DOI:10.2991/iccsee.2013.151.
11. M. Fateh, M. Rezvani, Y. Irani, A New Method of Coding for Steganography Based on LSB Matching Revisited, Security and Communication Networks (2021). DOI:10.1155/2021/6610678.
12. K. Tasdemir, F. Kurugollu, S. Sezer, Video steganalysis of LSB based motion vector steganography, in: European Workshop on Visual Information Processing (EUVIP), Paris, France, 2013, pp. 260-264.
13. K. Wu, Research of Video Steganalysis Algorithm Based on H265 Protocol, in: MATEC Web of Conferences 25, 03003 (2015). DOI:10.1051/matecconf/20152503003.
14. M. Fan, P. Liu, H. Wang, et al., Cross correlation feature mining for steganalysis of hash based least significant bit substitution video steganography, Telecommun Syst 63 (2016) 523–529. DOI:10.1007/s11235-016-0139-5.
15. A. Kobozieva, I. Bobok. Development of a New Approach to the Problem of Detecting the Integrity Violations of a Digital Image. Problemele Energeticii Regionale, 3 (59) (2023)109-121.
16. C. Bergman, J. Davidson, Unitary embedding for data hiding with the SVD, 2205. Online: https://dr.lib.iastate.edu/entities/publication/bb2b5041-1c92-4ff5-b7f4-ff73c3483eed.
17. T. Gloe, R. Bцhme, The “Dresden Image Database” for benchmarking digital image forensics, 2010 ACM Symposium on Applied Computing (SAC ’10), New York, 2010. pp. 1585–1591.
18. S. Geetha, S. Sindhu, N. Kamaraj, Close Color Pair Signature ensemble Adaptive Threshold based Steganalsis for LSB Embedding in Digital Images, Trans. Data Privacy 1 (2009) 140–161.
19. R. Srinivas, S. Panda, Performance analysis of various filters for image noise removal in different noise environment. International Journal of Advanced Computer Research 3 (2013) 47–52.
20. T.D. Sairam, K. Boopathybagan, Computational intelligence-based steganalysis comparison for RCM-DWT and PVA-MOD methods. Automatika 60 3 (2019) 285–293.
21. S. Mushtaq, A.H. Mir, Image copy move forgery detection: A Review. International Journal of Future Generation Communication and Networking 11 2 (2018) 11–22.
22. K. Asghar, Z. Habib, M. Hussain, Copy-move and splicing image forgery detection and localization techniques: A review. Australian Journal of Forensic Sciences 49 3 (2017) 281–307.

UDC 004.056.5

STEGANALYSIS METHOD FOR DETECTING LSB EMBEDDING IN DIGITAL VIDEO, DIGITAL IMAGE SEQUENCE

Dr.Sci. Bobok I., Dr.Sci. Kobozieva A.

National University "Odessa Polytechnic", Ukraine
E-mail: onu_metal@ukr.net; alla_kobozeva@ukr.net

Abstract. Steganography currently remains one of the most rapidly and effectively developing areas of information security. It leads to an increase in the relevance of steganalysis to avoid the negative consequences of covert communication for the purpose of illegal actions. Preference in the organization of a covert communication channel today is given to digital video. However, modern video steganalysis is behind the required level, particularly in conditions of low capacity of the covert channel. This paper is dedicated to developing a new effective steganalysis method for detecting embedded additional information in digital video (a sequence of digital images) by one of the most common steganographic methods today, which is the LSB method. The developed method is based on the sensitivity of linearity of singular vectors frequency in the corresponding original content matrices to disturbing influences. The authors determined this property in previous works. The current work established qualitative and quantitative differences between the properties of function, which represents the dependency of the singular vector frequency in the image/video frame matrix on its number for original contents and steganographic messages, which made it possible to develop an algorithmic implementation of the proposed method. The efficiency of this implementation in the case of a low capacity of a covert communication channel (<0.25 bit/pixel), including at a low frame embedding rate (<50%), exceeds the efficiency of analogues, which made it possible to increase the efficiency of steganalysis of digital video (sequence of digital images) as a whole. It is shown that the proposed method, after certain adaptation, can be used as a general expert method for detecting violations of the integrity of digital video/sequence of digital images under conditions of disturbing actions, in particular minor ones, which differ from steganomorphism.

Keywords: digital video, digital image, LSB method, steganalysis, covert channel capacity.

Section 4. Data Mining Technologies and Big Data

UDC 004.853

HELPFUL TELESCOPE TOOL FOR ASTRONOMICAL METADATA MINING FROM FITS FILES

Ph.D. Khlamov S.
Kharkiv National University of Radio Electronics, Ukraine
ORCID: 0000-0001-9434-1081
E-mail: sergii.khlamov@gmail.com

Abstract. *The chapter is devoted to the realization of the data mining approach related to the metadata of astronomical files from the big archives. Each astronomical file has the commonly defined structure, which contains the especial format of the metadata. Such metadata contain the necessary astronomical information, which is required for the proper storing, data mining, processing, analyzing under research. This realization was implemented as tool called "Telescope" using the C# programming language, .NET platform, Windows Forms technology and equipped with the MDB database file for Microsoft Access DBMS. The tool has two modes: console mode for the automated integration with the processing pipelines and mode with a graphical user interface (GUI) for the visualization of processing and the additional useful features. The Telescope tool was designed for mining the big astronomical data from the different archives, parsing the metadata from each astronomical file, and collecting it with the further insertion into the database. Such parsed data were used for the different purposes of the astronomical image processing and machine vision. The Telescope tool was developed during research under the CoLiTec project and was tested with the astronomical files from several archives on the different observatories. Also, the Telescope tool was successfully implemented and installed on the astronomical image processing pipelines in such observatories.*

Keywords: *data mining, big data, metadata, database, dataflow, image processing, machine vision, CoLiTec*

1. Introduction

Almost all astronomical frames are made by the CCD-camera [1] and can be received from the different sources: archives, servers, predefined series of frames, Virtual Observatories [2], clusters, etc. Each software for preparing the astronomical frames creates them as digital files in the FITS (Flexible Image Transport System) format [3]. This format is a digital files format for storing and transferring of their image and metadata (spreadsheets). Metadata is a kind of data, which provides the information about other data, except the original data content. There are a lot of different types of metadata, as following [4]: • Descriptive metadata is an information about the resource, which is used for identification and includes the elements, like author, title, abstract, and keywords. • Structural metadata is an information about the data containers and how the objects are collected. It includes the elements, like relationships, versions, types, etc. • Administrative metadata is an information for managing resources (creation date, edition date, permissions, etc.). • Reference metadata is an information about the static data, references to them, and contents. • Statistical metadata is an information about the processes for collection, producing, and publishing the statistical data. • Legal metadata is a legal information about the copyright, creator, and licensing.

The main purpose of metadata is to provide an information about the different aspects of original data and to summarize a basic information about it, which make tracking and processing it easier. The examples of metadata are as following: time and date of data creation, its meaning and purpose, creator or author, location, file size, used standards, sources, quality, etc. For example, the digital image includes the metadata, which describes image size, its color depth, resolution, creation time and date, exposure time, etc. A metadata of the text document includes an information about author, processing time of document, short summary, etc. The web pages include metadata, which describes a description of page content, and keywords linked to it. In astronomy metadata is used for the different image processing and machine vision purposes [5], like analyzing, acquiring, pre-processing, processing, and extraction of high-dimensional astronomical information [6]. Such purposes are focused on but not limited to the following tasks: filtering [7], brightness equalization [8] and background alignment [9], object's images detection [10], moving objects detection [11], astrometry of object's image [12], photometry of object's image [13], the estimation of the object's image and motion parameters [14], reference objects cataloging [15], objects recognition [16], matched filtration [17], time series analysis [18], Wavelet coherence analysis [19] and others. There are different types of astrophysical objects that can be recognized and detected, like galaxies, stars, robots [20, 21], drones [22], rockets, satellites [23], and even comets or asteroids [24]. Along with the software processing pipeline the data pipeline, known as a dataflow, is also performed. And when the processing is performed with the big astronomical data along with the metadata, the data mining approach is very useful [25]. The data mining is an analysis step of the "knowledge discovery in databases" (KDD) process [26]. The data mining carries out about the useful information extracting using the intelligent methods from a data set of metadata to transform it according to the required contracts and protocols and prepare for the further usage in the processing pipeline [27]. In this chapter we presented a description of the astronomical metadata from the real examples of CCD-images [28], usage in the data mining approach from the big archives, and its implementation as a developed Telescope tool, which is designed for mining the big astronomical data from the different archives, parsing the metadata from each astronomical file, and collecting it with the further insertion into the database. We selected one of the astronomical scientific software based on such processing pipeline, described the implementation of data mining purposes related to the astronomical processing metadata during invocation the image processing pipeline and its implementation in the developed Telescope tool. It is specially designed as a part of the CoLiTec project [29] for working with a big number of astronomical metadata that are used by the different mathematical and processing modules and components. Thus, the main aim of this chapter is the development and research of the helpful tool for the astronomical metadata mining. The rest of this chapter is organized as follows. Section 2 presents details about the big astronomical metadata. In section 3, the authors discuss in detail the processing pipeline of the selected scientific astronomical software called “CoLiTec”. In section 4, the authors discuss in detail the developed Telescope tool for mining the big astronomical data from the different storages and archives, parsing the metadata from each astronomical file, and collecting it with the further insertion into the database. Section 5 outlines the real astronomical examples and successful implementation of the mentioned helpful tool. The chapter ends with a conclusion in section 6 and acknowledgements in section 7.

2. Big astronomical metadata

Such a big volume of observational and historical astronomical information requires a lot of free storage space in hard disks, servers, clusters, etc. All this astronomical data can’t be collected on one single server, which will be universal and public for the scientific society. So, the disturbed systems, online catalogs with public access, cloud technologies should be used to resolve this complicated issue of storing astronomical information. The Virtual Observatory (VO) concept becomes more and more popular and important in astronomy, as a platform for easy access to the astronomical information stored in the different digital catalogs, archives, and servers [30]. The International Virtual Observatory Alliance (IVOA) has provided the different VO's standards and protocols for the realization and development of the VO. Also, IVOA has prepared special methods and algorithms for data interoperability. On a regular basis, the IVOA organizes the needed collaboration between the scientific society, information providers, and researchers. There are a lot of astronomical scientific programs, communities, and data centers that create and maintain the different image archives, catalogs with common astronomical information for KDD processing and data mining. For example, the SIMBAD database contains scientific data about more than 3 million astronomical objects [31]. It is based on the identifiers of 7.5 million astronomical objects. The SIMBAD database is continuously updated by the collaborative effort between all astronomical communities. The DAME (DAta Mining & Exploration) research contains a set of various web services that perform a scientific analysis of a big volume of astronomical information [32]. The KDD, data mining, and machine vision tasks for data analysis, gathering, cross-matching [33] and visualization from the big astronomical catalogs become more and more difficult. Because such big catalogs contain billions of astronomical objects and are continuously updated with new scientific data from fresh massive data sets, such as WISE (Wide-field Infrared Survey Explorer) [34], 2MASS (Two Micron All Sky Survey) [35], ESA Euclid space mission [36], ESA GAIA (Global Astrometric Interferometer for Astrophysics) space mission [37]. Collaboration of the new surveys projects and development of the new networks of the automated ground-based or space-based observational systems equipped with different large CCD cameras with a wide field of view (FOV) lead to a fast growth of the scientific information. Some of them are Large Synoptic Survey Telescope (LSST) [38], Pan-STARRS (Panoramic Survey Telescope and Rapid Response System) [39], and Thirty Meter Telescope (TMT) [40]. For example, the Pan-STARRS currently consists of two 1.8-m aperture telescopes (see Fig. 1). Each such telescope has a FOV of 3 square angular degrees and is equipped with the largest in the world CCD camera, which produces ~1.400 million pixels per astronomical image. And each such image requires ~2 Gb of free disk storage and the exposure time of one shot can be up to one minute.

Fig. 1. Panoramic Survey Telescope and Rapid Response System

The time for saving such astronomical images to disk storage is about one minute or even more. Totally, more than 10 Tb of scientific information is obtained every observational night and such data are taken as an online stream on a continuous basis. The LSST is currently under development. It is a reflecting telescope with an 8.4 meters primary mirror, which is a part of a wide-field survey. The design of each telescope includes 3 mirrors. Both have a very wide FOV of 3.5 square angular degrees. The resolution of the equipped CCD camera is 3.2 Gpixel (see Fig. 2). In the future LSST plans to take astronomical images every few nights by observing the whole sky. In such case, according to the planned LSST productivity, there will be more than 200 thousand uncompressed astronomical images per year, which will be more than 1.3 petabytes of scientific information. It will be a big challenge for LSST to manage the data mine, and processing of such big astronomical data in an effective way. The approximate requirements for servers and clusters for LSST are about 100 Tflops of power and about 15 Pb of disk storage. The TMT is currently under development. It will be an extremely large telescope (ELT), which is equipped with the segmented Ritchey-Chrétien telescope with a 492-segment primary mirror with a diameter of 30 meters (see Fig. 3).

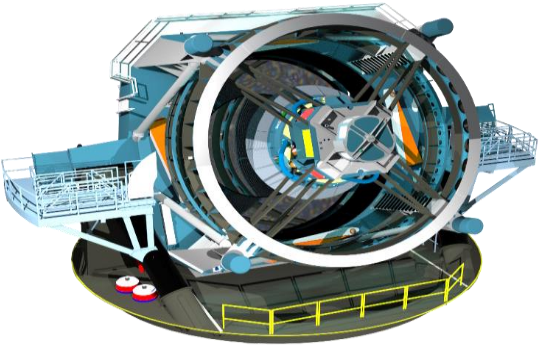


Fig. 2. Large Synoptic Survey Telescope

Fig. 3. Thirty Meter Telescope

The secondary mirror of TMT will have a diameter of 3.1 meters. The design of such telescope is adapted for the near-ultraviolet to mid-infrared (0.31 to 28 μm wavelengths) observations. Also, it will help in correcting image blur because of the adaptive optics. The TMT will be at the highest altitude of all the proposed ELTs, and it is supported on the government level by several nations. In the future TMT plans to collect each night about 90 Tb of the scientific astronomical data. There are a lot more such big telescopes, which produce a big volume of very huge astronomical images, like the Very Large Telescope (VLT) [41], the Canada-France-Hawaii Telescope (CFHT) in Hawaii, which is equipped with the MegaCam camera with a resolution of 16000 x 16000 (32 bits per pixel) [42]. The Sloan Digital Sky Survey (SDSS) is the most successful sky survey in astronomical history. The SDSS includes the most detailed Universe 3D maps with deep multi-color images of the third part of the whole sky, and spectra for more than 3 million astronomical objects [43]. So, in multi-directional astronomy the main goal is to collect and merge such big scientific data and astronomical information received from the big telescopes and surveys for further processing, which uses different technologies for data mining, machine vision, short time series analysis, and even Wavelet coherence analysis with the possibility of the further forecasting [44]. Almost all astronomical frames have a FITS format with standardized file structure and extension. In the common case the astronomical file extensions are: *.fits, *.FITS, *.fts, *.FTS, *.fit, *.FIT. Such FITS format is commonly used for the transformation, transferring, and archiving of astronomical data. The FITS format was developed by National Aeronautics and Space Administration (NASA) and is accepted as an international astronomical standard and is used by many astronomical and scientific organizations, like International Astronomical Union (IAU) [45], and other national and international organizations that deal with the astronomy or related scientific fields. The FITS format is commonly used for the storing the data without the image, like spectrums, photons list, data cubes or even structured data, such as databases with multiple tables. The FITS format includes many provisions to describe the photometric and spatial calibration, as well as image metadata. The structure of the FITS file consists of a header with metadata and a binary image. The header size is 2880 bytes and contain the list of human readable metadata in fixed string form of 80 symbols. Each string is an ASCII [46] stroke, which contains the pair with key and value, and have the common form: “KEYNAME = value / comment string”. Each header block should end with the especial key “END” with the empty value. The example of a header with metadata of the real astronomical FITS file is presented in Fig. 4. There are the minimum list of the required keywords to make the header and the whole FITS file valid. They are: “SIMPLE” (file conforms to FITS standard); “BITPIX” bitrade of FITS file, bits per pixel); “NAXIS” (number of axes); “NAXIS1” (number of points along axe 1); “NAXIS2” (number of points along axe 2); “END”.

```
SIMPLE  =                    T
BITPIX  =                   16 /8 unsigned int, 16 & 32 int, -32 & -64 real
NAXIS   =                    2 /number of axes
NAXIS1  =                  512 /fastest changing axis
NAXIS2  =                  512 /next to fastest changing axis
BSCALE  =   1.0000000000000000 /physical = BZERO + BSCALE*array_value
BZERO   =   32768.000000000000 /physical = BZERO + BSCALE*array_value
DATE-OBS= '2017-03-30T21:39:16' /YYYY-MM-DDThh:mm:ss observation start, UT
EXPTIME =   60.000000000000000 /Exposure time in seconds
EXPOSURE=   60.000000000000000 /Exposure time in seconds
SET-TEMP=  -40.000000000000000 /CCD temperature setpoint in C
CCD-TEMP=  -40.062500000000000 /CCD temperature at start of exposure in C
XPIXSZ  =   48.000000000000000 /Pixel Width in microns (after binning)
YPIXSZ  =   48.000000000000000 /Pixel Height in microns (after binning)
XBINNING=                    2 /Binning factor in width
YBINNING=                    2 /Binning factor in height
XORGSUBF=                    0 /Subframe X position in binned pixels
YORGSUBF=                    0 /Subframe Y position in binned pixels
READOUTM= '1 MPPS  ' /           Readout mode of image
FILTER  = 'V       ' /           Filter used when taking image
IMAGETYP= 'Light Frame' /        Type of image
OBJECT  = 'DO Dra  '
OBJCTRA = '11 43 38' /           Nominal Right Ascension of center of image
OBJCTDEC= '+71 41 20' /          Nominal Declination of center of image
OBJCTALT= ' 67.3413' /           Nominal altitude of center of image
OBJCTAZ = '  0.2784' /           Nominal azimuth of center of image
OBJCTHA = ' -0.0226' /           Nominal hour angle of center of image
SITELAT = '48 56 06' /           Latitude of the imaging location
SITELONG= '22 16 27' /           Longitude of the imaging location
JD      =   2457843.4022685187 /Julian Date at start of exposure
JD-HELIO=   2457843.4039910869 /Heliocentric Julian Date at exposure midpoint
AIRMASS =   1.0833536254388081 /Relative optical path length through atmosphere
FOCALLEN=   9000.0000000000000 /Focal length of telescope in mm
APTDIA  =   1000.0000000000000 /Aperture diameter of telescope in mm
APTAREA =   777544.20340061188 /Aperture area of telescope in mm^2
SWCREATE= 'MaxIm DL Version 5.12' /Name of software that created the image
SBSTDVER= 'SBFITSEXT Version 1.0' /Version of SBFITSEXT standard in effect
TELESCOP= 'VNT     ' /           telescope used to acquire this image
INSTRUME= 'FLI     ' /           instrument or camera used
OBSERVER= 'DPV     '
NOTES   = '        '
FLIPSTAT= '        '
SWOWNER = 'Amigo   ' /           Licensed owner of software
END
```

Fig. 4. The example of a header with metadata of the real astronomical FITS file

3. Processing pipeline in the CoLiTec software

For our research we have selected astronomical scientific software for detection of the moving objects in a series of CCD-frames called "CoLiTec" (Collection Light Technology), which implements the image processing pipeline. Such software performs almost all astronomical image processing tasks, like filtering [47], brightness equalization [48], background alignment, image stacking/segmentation [49], object detection, motion detection [14], object astrometry [50], object photometry [8], object's image and motion parameters estimation [51], machine (computer) vision [52] of the reference objects to be cataloged [15], object recognition [53], time series analysis [18], Wavelet coherence analysis [54], machine learning recognition [16] and others. The modern CoLiTec software was developed using different technologies and approaches for big data processing, data mining, and machine vision. In the astronomy direction, the CoLiTec software is designed to perform the following main stages of machine vision and image processing: pre-processing (astronomical data collection -> worst data rejection -> useful data extraction -> classification -> clustering -> background alignment -> brightness equalization), processing (recognition patterns applying -> machine vision -> object's image detection -> astrometry -> photometry -> tracks detection), knowledge extraction (astronomical objects to be discovered, tracks parameters for the investigation, light curves of the variable stars). More features of the CoLiTec software are detailed described below [29]:

– processing images with the very wide FOV (<10 degrees^2);

– calibration and cosmetic correction in automated mode using the appropriate calibration master frames and their creation if necessary;

– brightness equalization and background alignment of the images in series using the mathematical inverse median filter;

– rejection of the bad and unclear observations and measurements of the investigated astronomical objects in automated mode;

– fully automatic robust algorithm of the astrometric and photometric reduction of the investigated astronomical objects;

– detection of the faint investigated astronomical moving objects in a series of CCD images with a signal-to-noise ratio (SNR) of more than 2.5 in automated mode;

– detection of the very fast investigated astronomical objects (<40.0 pix./frame) in automated mode;

– detection of the astronomical objects with near-zero apparent motion (from 0.7 pix./frame) in automated mode;

– rejection of the investigated astronomical objects with bad and corrupted measurements in automated mode;

– viewer of the processing results with simple and understandable graphical user interface (GUI) by the LookSky software;

– confirmation of the most interesting astronomical objects at the night of their preliminary discovery;

– multi-threaded processing support;

– multi-cores systems support;

– support of managing the individual threads and processes;

– processing pipeline managed by the On-line Data Analysis System (OLDAS);

– deciding system of the processing results, which allows to adapt the end-user settings and inform the user about the processing results at each stage in the pipeline;

– subject mediator as a major part of the data control in the pipeline during processing.

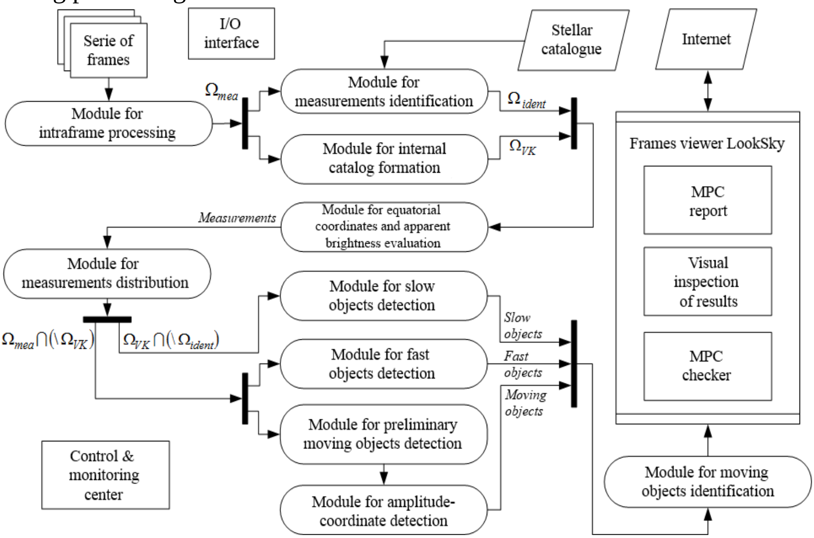


Fig. 5. The high level processing pipeline of the CoLiTec software

CoLiTec software realizes the different knowledge discovery in databases and data mining approaches, like pre-processing, clustering, classification, identification, processing, summarization.

CoLiTec software is a very complex astronomical system for the big data sets processing, which includes the different features, user-friendly tools for the processing management, results reviewing, integration with online astronomical catalogs and a lot of computational components and modules that are based on the developed mathematical methods [10, 14, 51].

The high level processing pipeline with the developed modules and implemented methods of the CoLiTec software is presented in Fig. 5.

4. Metadata mining by the Telescope tool

Under the research in scope of the CoLiTec project [29] we have developed the Telescope tool for mining the big astronomical data from the different storages and archives, parsing the metadata from each astronomical file, and collecting it with the further insertion into the database. Parsed data were used for the different purposes of astronomical image processing and machine vision. The Telescope tool realized the different data mining tasks, like receiving, storing, selecting, preprocessing, transforming, useful data extraction, classification, and knowledge discovery in databases (KDD) [27].

As a database the developers have selected a MDB database file as a native format for the Microsoft Access DBMS with ".mdb" extension [55]. There are different fields located in several tables of the MDB database file in the Telescope tool that represent the necessary for research metadata of the astronomical file. The list of such fields with their description and types including the system fields are presented in Table 1. The following stack of technologies were used for the software development: C# programming language, .NET platform, Windows Forms technology and MDB database file for Microsoft Access database management system (DBMS). The Windows Forms (WinForms) is an open-source and free graphical library, which is in scope of the Microsoft .NET Framework [56] and play the role as a platform for developing the client applications for desktop, laptop, and tablet PCs.

The Telescope tool has two different modes: console mode for the silent automated integration with the image processing pipelines and mode with a graphical user interface (GUI) for the visualization of processing and the additional useful features. The console mode of the Telescope tool is designed for the integration with processing pipeline as a precondition step for searching, collecting, and parsing the astronomical metadata from the files in archives.

Table 2

Metadata in header of the FITS file

Field	Type	Description
FitCode	Long Text	Unique code of each FITS file
FitName	Short Text	Name of FITS file on storage
FitPath	Short Text	Full path to FITS file on storage
RA	Number	Right ascension of frame center
DE	Number	Declination of frame center
Date	Date	Date of observation
Time	Time	Time of observation
Exp	Number	Exposure time
BitPix	Number	Pixel bit rate
Height	Number	Height of frame
Width	Number	Width of frame
PixHeight	Number	Height of pixel
PixWidth	Number	Width of pixel
Focus	Number	Focal length of telescope
Apertrure	Number	Aperture size of telescope
Temp	Number	Temperature of CCD-camera
Long	Number	Longitude of telescope
Lat	Number	Latitude of telescope
Alt	Number	Altitude of telescope
Instrum	Short Text	Instrument name
Telescope	Short Text	Telescope name
Observer	Short Text	Observer name
Obj	Short Text	Investigated object name

The Telescope tool in console mode can be launched in the Windows Command Prompt using the following list of the implemented commands for: showing help information with all available commands (“help”); adding the new astronomical FITS file with metadata to be parsed and inserted into the database (“add [PathToFile.fits]”); adding the new already parsed astronomical metadata from the text file into the database (“add [PathToFile.txt]”); searching for the metadata by the different optional criteria in the database (“find [Observer] [Telescope] [RA] [DE] [Date] [Time] [PathToFile.txt]”); searching for the metadata in a range by the especial criteria in the database (“find [RAfrom] [RAto] [DEfrom] [DEto] [PathToFile.txt]”); exporting the MDB database file to the user’s local folder ("export [PathToFolder]"); importing the MDB database file from the user's local folder ("import [PathToDBFile.mdb]").

The mode with GUI of the Telescope tool is designed for the independent big astronomical metadata preparation. The processing pipeline includes the following steps:

– selecting the work folder including subfolders with the different astronomical FITS files (see Fig. 6);

– recurrency searching for the astronomical FITS files according to the extensions in the work folder and forming the results list (see Fig. 7 (left));

– parsing the astronomical FITS files from the results list, extracting the metadata and insertion it into the MDB database file (see Fig. 7 (right));

– exporting/importing the MDB database file;

– logging the searching, parsing process, error handling;

– finding the different astronomical metadata in the MDB database file and forming the results list (see Fig. 8 (left));

– converting from arcseconds (angular hours / minutes / seconds) to decimal (2), (3) and vice (4), (5) (see Fig. 8 (right)).

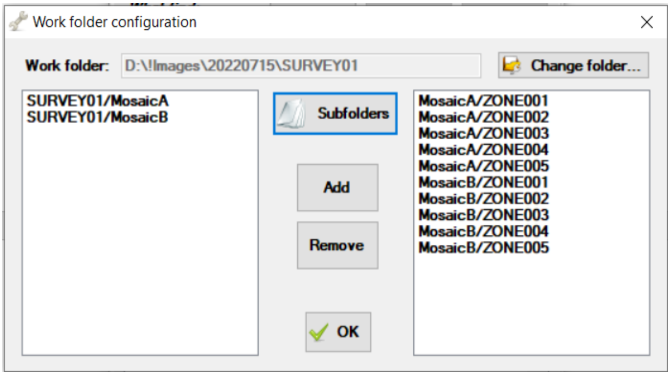


Fig.6. "Work folder configuration" window in the Telescope tool

In general, positional coordinates of objects can be of two types: decimal Cartesian (x and y) in the image plane and stellar (right ascension RA and declination DE) as angular coordinates in the sky.

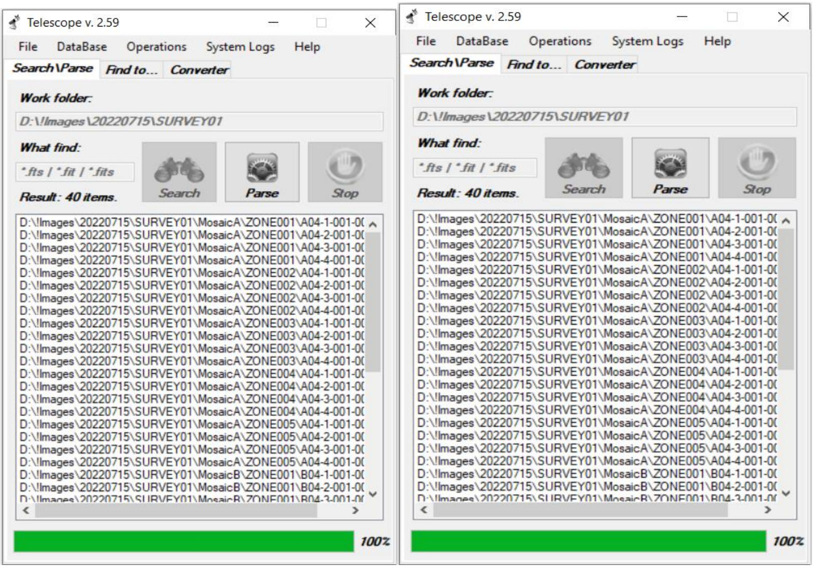


Fig. 7. The "Search/Parse" page with the list of results in the Telescope tool during searching procedure (left) and during parsing procedure (right)

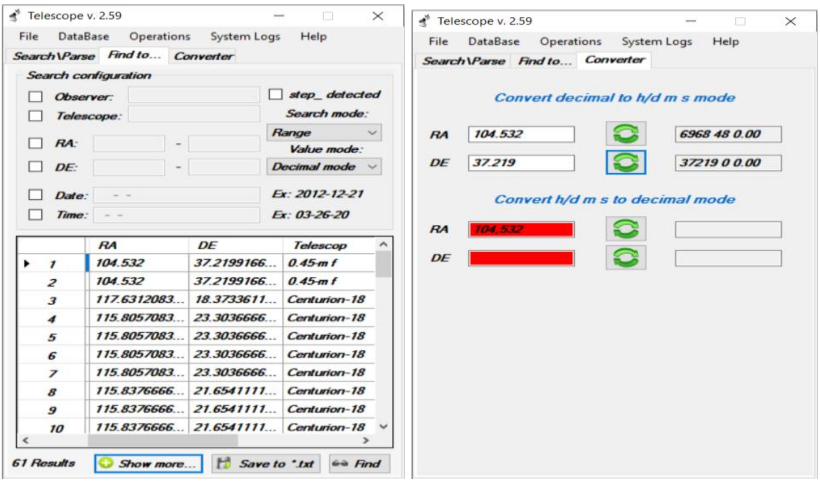


Fig.8. The "Find to" page (left) and the "Convertor" page (right) in the Telescope tool

So, to use the appropriate type of positional coordinates, the following mutual recalculation can be used:

$$x_d = \left(\left(\left(\frac{RA_S}{60} + RA_M\right)/60\right) + RA_H\right) * 15; \quad (2)$$

$$y_d = \left(\left(\frac{DE_S}{60} + DE_M\right)/60\right) + DE_H; \quad (3)$$

$$RA_H{:}RA_M{:}RA_S = \left\{\begin{matrix} (int)\dfrac{x_d}{15} \\ (int)(x_d - RA_H) * 60 \\ (x_d - RA_H - RA_M) * 60 \end{matrix}\right\}; \quad (4)$$

$$DE_H{:}DE_M{:}DE_S = \left\{\begin{matrix} (int)y_d \\ (int)(y_d - DE_H) * 60 \\ (y_d - DE_H - DE_M) * 60 \end{matrix}\right\}, \quad (5)$$

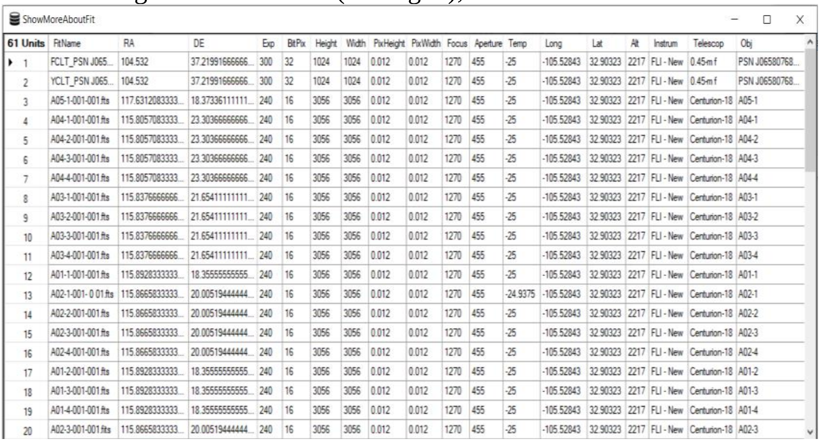
where *d* is a decimal value;

H are the angular hours;

M are the angular minutes;

S are the angular seconds.

– displaying all available astronomical metadata in the MDB database file according to the results list (see Fig. 9);



ShowMoreAboutFit

61 Units	FitName	RA	DE	Exp	BitPix	Height	Width	PixHeight	PixWidth	Focus	Aperture	Temp	Long	Lat	Alt	Instrum	Telescop	Obj
1	FCLT_PSN J065...	104.532	37.21991666666...	300	32	1024	1024	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	0.45-m f	PSN J06580768...
2	YCLT_PSN J065...	104.532	37.21991666666...	300	32	1024	1024	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	0.45-m f	PSN J06580768...
3	A05-1-001-001.fts	117.6312083333...	18.37336111111...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A05-1
4	A04-1-001-001.fts	115.8057083333...	23.30366666666...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A04-1
5	A04-2-001-001.fts	115.8057083333...	23.30366666666...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A04-2
6	A04-3-001-001.fts	115.8057083333...	23.30366666666...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A04-3
7	A04-4-001-001.fts	115.8057083333...	23.30366666666...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A04-4
8	A03-1-001-001.fts	115.8376666666...	21.65411111111...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A03-1
9	A03-2-001-001.fts	115.8376666666...	21.65411111111...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A03-2
10	A03-3-001-001.fts	115.8376666666...	21.65411111111...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A03-3
11	A03-4-001-001.fts	115.8376666666...	21.65411111111...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A03-4
12	A01-1-001-001.fts	115.8928333333...	18.35555555555...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A01-1
13	A02-1-001- 0 01.fts	115.8665833333...	20.00519444444...	240	16	3056	3056	0.012	0.012	1270	455	-24.9375	-105.52843	32.90323	2217	FLI - New	Centurion-18	A02-1
14	A02-2-001-001.fts	115.8665833333...	20.00519444444...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A02-2
15	A02-3-001-001.fts	115.8665833333...	20.00519444444...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A02-3
16	A02-4-001-001.fts	115.8665833333...	20.00519444444...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A02-4
17	A01-2-001-001.fts	115.8928333333...	18.35555555555...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A01-2
18	A01-3-001-001.fts	115.8928333333...	18.35555555555...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A01-3
19	A01-4-001-001.fts	115.8928333333...	18.35555555555...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A01-4
20	A02-3-001-001.fits	115.8665833333...	20.00519444444...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A02-3

Fig. 9. Astronomical metadata in the database according to the results list in the Telescope tool

The relationships between tables and fields in the Microsoft Access solution of the MDB database file integrated into the Telescope tool is presented in Fig. 10. The example of the filled “Fits” table in the MDB database file integrated into the Telescope tool with the real astronomical data is presented in Fig. 11.

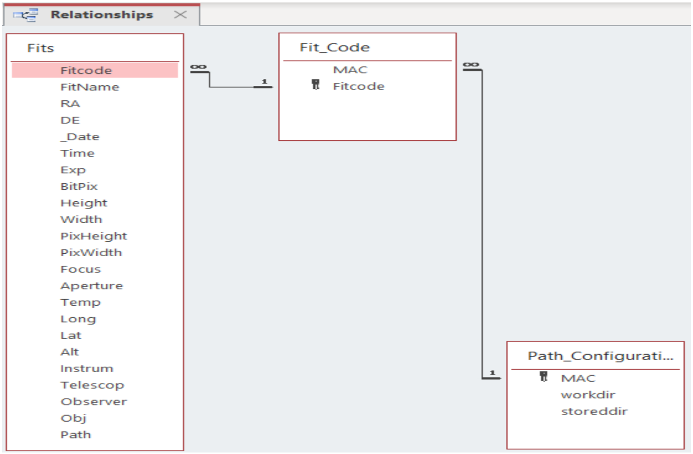


Fig. 10. Relationships between tables and fields in the in the Microsoft Access solution integrated into the Telescope tool

Fits

Fitcode	FitName	RA	DE	Time	Exp	BitPix	Height	Width	PixHeight	PixWidth	Focus	Long	Lat
B04-3-001-001_Centurion-18_	B04-3-001-001.fts	179.175291666667	-0.824583333333333	10-38-21	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
B04-4-001-001_Centurion-18_	B04-4-001-001.fts	179.175291666667	-0.824583333333333	10-58-42	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A04-3-001-001_Centurion-18_	A04-3-001-001.fts	115.892833333333	18.3555555555556	02-47-17	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A04-4-001-001_Centurion-18_	A04-4-001-001.fts	115.892833333333	18.3555555555556	03-06-46	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A01-2-001-001_Centurion-18_	A01-2-001-001.fts	115.892833333333	18.3555555555556	03-06-46	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A01-3-001-001_Centurion-18_	A01-3-001-001.fts	115.892833333333	18.3555555555556	03-26-20	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A03-1-001-001_Centurion-18_	A03-1-001-001.fts	115.892833333333	18.3555555555556	03-26-20	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A03-2-001-001_Centurion-18_	A03-2-001-001.fts	115.892833333333	18.3555555555556	03-45-51	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A01-4-001-001_Centurion-18_	A01-4-001-001.fts	115.892833333333	18.3555555555556	03-45-51	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A01-1-001-001_Centurion-18_	A01-1-001-001.fts	115.892833333333	18.3555555555556	02-47-17	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A02-3-001-001s_Centurion-18	A02-3-001-001.fits	115.892833333333	18.3555555555556	02-47-17	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
FCLT_PSN J06580768+371311	FCLT_PSN J065807	115.892833333333	18.3555555555556	02-47-17	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
YCLT_PSN J06580768+371311	YCLT_PSN J065807	115.866583333333	20.0051944444444	02-52-07	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
FCLT_PSN J06580768+371311	FCLT_PSN J065807	115.866583333333	20.0051944444444	02-52-07	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A02-1-001- 0 01_Centurion-18	A02-1-001- 0 01.ft	115.866583333333	20.0051944444444	02-52-07	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A05-1-001-001_Centurion-18_	A05-1-001-001.fts	115.866583333333	20.0051944444444	03-11-37	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
YCLT_PSN J06580768+371311	YCLT_PSN J065807	115.866583333333	20.0051944444444	03-11-37	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A02-2-001-001_Centurion-18_	A02-2-001-001.fts	115.866583333333	20.0051944444444	03-11-37	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A04-4-001-001_Centurion-18_	A04-4-001-001.fts	115.866583333333	20.0051944444444	03-31-11	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A03-4-001-001_Centurion-18_	A03-4-001-001.fts	115.866583333333	20.0051944444444	03-31-11	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A03-1-001-001_Centurion-18_	A03-1-001-001.fts	115.866583333333	20.0051944444444	03-31-11	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A02-3-001-001s_Centurion-18	A02-3-001-001.fits	115.866583333333	20.0051944444444	03-31-11	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323

Fig. 11. Astronomical metadata in the database according to the results list in the Telescope tool

Each value from metadata in header of the FITS file found by the Telescope tool was successfully parsed and filled into the appropriate field in the "Fits" table in the MDB database file for further using and processing.

All interactions between such a database file and the Telescope tool are realized by using the structured query language (SQL) as a programming language for storing and processing information in a relational database. The

Telescope tool implements the following algorithm for astronomical metadata mining.

1. Selecting the work folder including subfolders with the different astronomical FITS or TXT files.
2. Database import from the pre-filled MDB database file, which was previously exported from the last session.
3. Recurrency searching for the astronomical FITS or TXT files according to the extensions in the work folder.
4. Getting access to the astronomical FITS or TXT files (in case when location of big astronomical data is in the different remote/web archives) and download them.
5. Reading the astronomical FITS or TXT files (for FITS files splitting for two parts: header with astronomical information and body with image bytes).
6. Parsing the astronomical FITS or TXT files.
7. Astronomical data receiving from the astronomical FITS or TXT files (for FITS files from header).
8. Astronomical data converting using the different mathematical methods (e.g., positional coordinates conversion).
9. Astronomical data structurization according to the patterns/classes/clusters based on the statistical modeling [57].
10. Metadata mining from the astronomical data structure according to the appropriate parameters/fields/properties.
11. Filling in the appropriate fields and tables in database using the metadata.
12. Database export to the MDB database file (if needed).

The metadata mining algorithm implemented in the Telescope tool is presented as UML-diagram in Fig. 12.

5. Practical implementation with the real astronomical examples

The Telescope tool in scope of the CoLiTec software were installed in the different observatories (Mayaki Astronomical Observatory [58, 59], ISON-NM and ISON-Kislovodsk observatories, Vihorlat Observatory [8, 48]), astronomical archives [2], and Ukrainian Virtual Observatory (UkrVO) [30]. More detailed information about the observatories, telescopes equipped by the different CCD-cameras as well as their parameters are provided below. The Mayaki observing station of "Astronomical Observatory" Research Institute of I. I. Mechnikov Odessa National University has the 0.48 m AZT-3 telescope – reflector with focal length 2025 mm and CCD-camera Sony ICX429ALL (resolution 795×596).

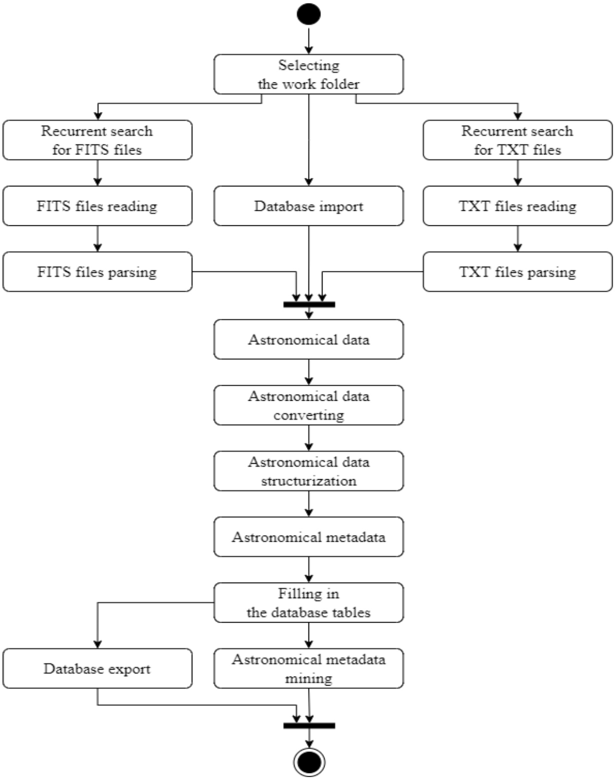


Fig. 12. Metadata mining algorithm implemented in the Telescope tool

The observatory "ISON-NM observatory" has the 0.4 m SANTEL-400AN telescope with CCD-camera FLI ML09000-65 (3056×3056 pixels, 12 microns). The observatory "ISON-Kislovodsk" has the 19.2 cm wide-field GENON (VT-78) telescope with CCD-camera FLI ML09000-65 (4008×2672 pixels, 9 microns). The observatory "Vihorlat Observatory in Humenne" has the Vihorlat National Telescope (VNT) – Kassegren telescope with 1 m main mirror with focal length 8925 mm and CCD-camera FLI PL1001E (512×512 pixels). The Vihorlat Observatory also has the Celestron C11 telescope – Schmidt-Cassegrain telescope with 28 cm main mirror with focal length 3060 mm and CCD-camera G2-1600 (resolution 768×512 pixels). The data mining of astronomical metadata by the Telescope tool was performed during the processing of up to 1 million astronomical files both archived and original formed from the different telescopes. Such set of data was processed by the Telescope tool based on the metadata in their headers. An example of metadata mining of the of astronomical files by the Telescope tool integrated into the processing pipeline is presented in Fig. 13.

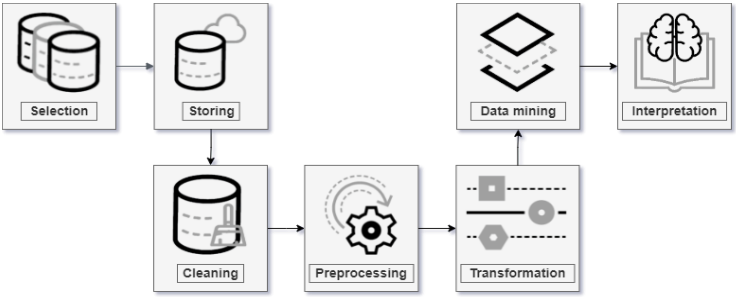


Fig. 13. Example of metadata mining of the of astronomical files by the Telescope tool integrated into the processing pipeline

The received information was inserted into the database as a big data and processed by the UkrVO, which also processed a lot of different big astronomical archives both digital and even plates.

6. Conclusion

The Telescope tool with the realization of the data mining approach related to the metadata of astronomical files from the big archives was developed. The Telescope tool is implemented using the C# programming language, .NET platform, Windows Forms technology and equipped with the MDB database file for the Microsoft Access DBMS. The SQL was used as a programming language for storing and processing information in a relational database. The tool has two modes: console mode for the automated integration with the processing pipelines and mode with a graphical user interface (GUI) for the visualization of processing and the additional useful features. The Telescope tool was designed for mining the big astronomical data from the different archives, parsing the metadata from each astronomical file, and collecting it with the further insertion into the database. Parsed data were used for the different purposes of astronomical image processing and even for the Wavelet coherence analysis purposes. The Telescope tool was also developed during research under the CoLiTec project. It was tested with up to 1 million astronomical files from several archives on the different observatories. Such archives included astronomical files of different formats and types of metadata. All such metadata was parsed and structured, which given us an opportunity to perform the proper metadata mining. Such proper metadata was used in the further research and calculations, where the measurements of each known objects are used to clarify the typical form of image, its orbits, motion parameters and other important astronomical properties on the long historical period. Also, the Telescope tool was successfully implemented and installed on the astronomical image processing pipelines in such observatories.

7. Acknowledgements

The authors thank all observatories for their collaboration, providing an access to their archives with astronomical data and possibility to install the Telescope tool in scope of the CoLiTec software as a main astronomical processing pipeline to conduct the current research and test developed tool.

8. References

[1] H.F. Zhang, C. Chen, J. Wang, et al., Scientific CCD camera for the CSTAR2 telescope in Antarctica, Journal of Astronomical Telescopes, Instruments, and Systems 5 3 (2019). DOI: 10.1117/1.JATIS.5.3.036002.

[2] I.B. Vavilova, et al., Surveys, catalogues, databases, and archives of astronomical data, Knowledge Discovery in Big Data from Astronomy and Earth Observation, Astrogeoinformatics (2020) 57-102. DOI: 10.1016/B978-0-12-819154-5.00015-1.

[3] W.D. Pence, L. Chiappetti, C.G. Page, R.A. Shaw, E. Stobie, Definition of the flexible image transport system (fits), version 3.0, Astronomy & Astrophysics 524 42 (2010). DOI: 10.1051/0004-6361/201015362.

[4] J. Furner, Definitions of “metadata”: A brief survey of international standards, Journal of the Association for Information Science and Technology, 71 6 (2020) 33-42. DOI: 10.1002/asi.24295.

[5] C. Steger, M. Ulrich, C. Wiedemann, Machine vision algorithms and applications, John Wiley & Sons, 2018.

[6] V. Akhmetov, et al., New approach for pixelization of big astronomical data for machine vision purpose, in: IEEE International Symposium on Industrial Electronics, 2019, pp. 1706–1710. DOI:10.1109/ISIE.2019.8781270.

[7] C.A. Politsch, et al., Trend filtering–I. A modern statistical tool for time-domain astronomy and astronomical spectroscopy, in: MNRAS, vol. 492, issue 3, 2020, pp. 4005-4018. DOI: 10.1093/mnras/staa106.
[8] Š. Parimucha, et al., CoLiTecVS – A new tool for an automated reduction of photometric observations, Contributions of the Astronomical Observatory Skalnate Pleso, vol. 49, issue 2, 2019, pp. 151-153.
[9] E. Corsaro, et al., Spin alignment of stars in old open clusters, Nature Astronomy 1 4 (2017) 1-6. DOI:10.1038/s41550-017-0064.
[10] L. Mykhailova, et al., Method of maximum likelihood estimation of compact group objects location on CCD-frame, Eastern-European Journal of Enterprise Technologies 5 4 (2014) 16-22. DOI:10.15587/1729-4061.2014.28028.
[11] S. Khlamov, et al., Development of computational method for detection of the object's near-zero apparent motion on the series of CCD–frames, Eastern-European Journal of Enterprise Technologies, 9 80 (2016) 41-48. DOI:10.15587/1729-4061.2016.65999.
[12] V. Akhmetov, et al., Astrometric reduction of the wide-field images, Advances in Intelligent Systems and Computing 1080 (2020) 896–909. DOI:10.1007/978-3-030-33695-0_58.
[13] V. Savanevych, et al., CoLiTecVS software for the automated reduction of photometric observations in CCD-frames, Astronomy and Computing, 40 15 (2022). DOI: 10.1016/j.ascom.2022.100605.
[14] V. Savanevych, et al., A method of immediate detection of objects with a near-zero apparent motion in series of CCD-frames. Astronomy & Astrophysics, 609 11 (2018). DOI: 10.1051/0004-6361/201630323.
[15] V. Savanevych, et al., Selection of the reference stars for astrometric reduction of CCD-frames, Advances in Intelligent Systems and Computing 1080 (2020) 881–895. DOI: 10.1007/978-3-030-33695-0_57.
[16] L. Kirichenko, P. Zinchenko, T. Radivilova, Classification of time realizations using machine learning recognition of recurrence plots, Advances in Intelligent Systems and Computing 1246 (2021) 687–696. DOI: 10.1007/978-3-030-54215-3_44.
[17] S. Khlamov, et al., Development of computational method for matched filtration with analytic profile of the blurred digital image, Eastern-European Journal of Enterprise Technologies 5 119 (2022) 24–32. DOI: 10.15587/1729-4061.2022.265309.
[18] L. Kirichenko, et al., Generalized approach to analysis of multifractal properties from short time series, International Journal of Advanced Computer Science and Applications 11 5 (2020) 183–198. DOI: 10.14569/IJACSA.2020.0110527.
[19] V. Baranova, et al., Wavelet Coherence as a Tool for Studying of Economic Dynamics in Infocommunication Systems, in: IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T), 2019, pp. 336-340. DOI: 10.1109/PICST47496.2019.9061301.
[20] A. Tantsiura, et al., Evaluation of the potential accuracy of correlation extreme navigation systems of low-altitude mobile robots, International Journal of Advanced Trends in Computer Science and Engineering 8 5 (2019) 2161–2166. DOI: 10.30534/ijatcse/2019/47852019.
[21] N. Yeromina, et al., Method of reference image selection to provide high-speed aircraft navigation under conditions of rapid change of flight trajectory, International Journal of Advanced Technology and Engineering Exploration 8 85 (2021) 1621–1638. DOI: 10.19101/IJATEE.2021.874814.
[22] M. Ivanov, et al., Effective informational entropy reduction in multi-robot systems based on real-time TVS, in: IEEE International Symposium on Industrial Electronics, 2019, pp. 1162–1167. DOI: 10.1109/ISIE.2019.8781209.

[23] V. Akhmetov, et al., Cloud computing analysis of Indian ASAT test on March 27, 2019, Proceedings of the 2019 IEEE International Scientific-Practical Conference: Problems of Infocommunications Science and Technology, PIC S and T 2019, 2019, pp. 315–318. DOI: 10.1109/PICST47496.2019.9061243.
[24] D. Oszkiewicz, V. Troianskyi, et al., Spins and shapes of basaltic asteroids and the missing mantle problem. Icarus 397 (2023). DOI: 10.1016/j.icarus.2023.115520
[25] S. Khlamov, et al., Data Mining of the Astronomical Images by the CoLiTec Software, CEUR Workshop Proceedings, vol. 3171, 2022, pp. 1043–1055.
[26] Q. Luo, Advancing knowledge discovery and data mining, First International Workshop on Knowledge Discovery and Data Mining (WKDD 2008), IEEE, 2008, pp. 3-5. DOI: 10.1109/WKDD.2008.153.
[27] Y. Zhang, Y. Zhao, C. Cui, Data mining and knowledge discovery in database of astronomy, Progress in Astronomy 20 4 (2002) 312-323.
[28] G.K. Adam, et al., Embedded microcontroller with a CCD camera as a digital lighting control system, Electronics 8 1 (2019) 33. DOI:10.3390/electronics8010033.
[29] S. Khlamov, V. Savanevych, Big astronomical datasets and discovery of new celestial bodies in the Solar System in automated mode by the CoLiTec software, Knowledge Discovery in Big Data from Astronomy and Earth Observation, Astrogeoinformatics (2020) 331–345. DOI: 10.1016/B978-0-12-819154-5.00030-8.
[30] I. B. Vavilova, et al., Astrometry and photometry of asteroids from the UkrVO database of astroplates, in: Proceedings of the IAU, vol. 15, issue S364, 2021, pp. 239-245. DOI: 10.1017/S1743921322000047.
[31] L. Kuhn, M. Shubat, P. Barmby, Comparing NED and SIMBAD classifications across the contents of nearby galaxies. Monthly Notices of the Royal Astronomical Society, 515 1 (2022) 807-816. DOI: 10.1093/mnras/stac1801.
[32] S. Cavuoti, M. Brescia, G. Longo, Data mining and knowledge discovery resources for astronomy in the Web 2.0 age. SPIE Astronomical Telescopes and Instrumentation, Software and Cyberinfrastructure for Astronomy II 8451 (2012). DOI: 10.1117/12.925321.
[33] V. Akhmetov, S. Khlamov, A. Dmytrenko, Fast coordinate cross-match tool for large astronomical catalogue. Advances in Intelligent Systems and Computing 871 (2019) 3–16. DOI: 10.1007/978-3-030-01069-0_1.
[34] E. Wright, et al., The Wide-field Infrared Survey Explorer (WISE): Mission Description and Initial On-orbit Performance. AJ, vol. 140, 2010, pp. 1868-1881. DOI: 10.1088/0004-6256/140/6/1868.
[35] M., Skrutskie, et al., The two micron all sky survey (2MASS). The Astronomical Journal, 131 (2) (2006). DOI: 10.1086/498708.
[36] G. Racca, et al., The Euclid mission design. Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave, in: Proceedings of the SPIE, 9904, 2016, pp. 235-257. DOI: 10.1117/12.2230762.
[37] B. Carry, W. Thuillot, F. Spoto, et al., Potential asteroid discoveries by the ESA Gaia mission: Results from follow-up observations. Astronomy and Astrophysics, 648 A96 (2021). DOI: 10.1051/0004-6361/202039579.
[38] K. Borne, et al., The LSST data mining research agenda, in: AIP Conference Proceedings, 1082 (1), 2008, pp. 347-351. DOI: 10.1063/1.3059074.
[39] N. Kaiser, et al., Pan-STARRS: a large synoptic survey telescope array, in: Proceedings of the Conference on Survey and Other Telescope Technologies and Discoveries, vol. 4836, 2002, pp. 154-164. DOI: 10.1117/12.457365.

[40] J. Nelson, G. Sanders, The status of the Thirty Meter Telescope project. Ground-based and Airborne Telescopes II, 7012: 70121A, 18, International Society for Optics and Photonics, 2008. DOI: 10.1117/12.788238.
[41] J-L. Beuzit, et al., SPHERE: the exoplanet imager for the Very Large Telescope. Astronomy & Astrophysics, 631, A155, 2019. DOI: 10.1051/0004-6361/201935251.
[42] O. Ilbert, et al., Accurate photometric redshifts for the CFHT legacy survey calibrated using the VIMOS VLT deep survey. Astronomy & Astrophysics, 457 (3) (2006) 841-856. DOI: 10.1051/0004-6361:20065138.
[43] D. Eisenstein, et al., SDSS-III: Massive spectroscopic surveys of the distant universe, the Milky Way, and extra-solar planetary systems. The Astronomical Journal, 142 (3) (2011). DOI: 10.1088/0004-6256/142/3/72.
[44] S. Dombrovska, et al., The prospects of integration of the advertising market of Ukraine into the global advertising business. Problems and Perspectives in Management 16 (2) (2018) 321-330. DOI:10.21511/ppm.16(2).2018.29.
[45] J. Andersen, D. Baneke, C. Madsen, The International Astronomical Union: Uniting the Community for 100 Years, Springer, 2019.
[46] H. T. Elshoush, M. M. Mahmoud, A. Abdelrahman, A new high capacity and secure image realization steganography based on ASCII code matching, Multimedia Tools and Applications 81 4 (2022) 5191-5237. DOI:10.1007/s11042-021-11741-y.
[47] S. Khlamov, et al., Development of the matched filtration of a blurred digital image using its typical form. Eastern-European Journal of Enterprise Technologies 1 9 (2023) 62-71. DOI: 10.15587/1729-4061.2023.273674.
[48] I. Kudzej, et al., CoLiTecVS – A new tool for the automated reduction of photometric observations, Astronomische Nachrichten 340 1-3 (2019) 68-70. DOI:10.1002/asna.201913562.
[49] R. Gonzalez, R. Woods, Digital image processing, 4th edition, NY: Pearson, 2018.
[50] V. Savanevych, et al., Mathematical methods for an accurate navigation of the robotic telescopes, Mathematics 11 10 (2023). DOI:10.3390/math11102246.
[51] V. Savanevych, et al., Formation of a typical form of an object image in a series of digital frames, Eastern-European Journal of Enterprise Technologies 6 120 (2022) 51–59. DOI: 10.15587/1729-4061.2022.266988.
[52] R. Klette, Concise computer vision, Springer, London, 2014.
[53] S. Khlamov, I. Tabakova, T. Trunova, Recognition of the astronomical images using the Sobel filter, in: Proceedings of the International Conference on Systems, Signals, and Image Processing, 4, 2022. DOI: 10.1109/IWSSIP55020.2022.9854425.
[54] M. Dadkhah, et al., Methodology of wavelet analysis in research of dynamics of phishing attacks, International Journal of Advanced Intelligence Paradigms 12 (2019) 220-238. DOI: 10.1504/IJAIP.2019.098561.
[55] J. Eckstein, B. R. Schultz, Introductory relational database design for business, with Microsoft Access, John Wiley & Sons, 2018.
[56] S. Bisser, Introduction to the Microsoft Bot Framework, Microsoft Conversational AI Platform for Developers (2021) 25-66. DOI: 10.1007/978-1-4842-6837-7_2.
[57] V. Shvedun, S. Khlamov, Statistical modelling for determination of perspective number of advertising legislation violations, Actual Problems of Economics 184 10 (2016) 389-396.
[58] V. Troianskyi, P. Kankiewicz, D. Oszkiewicz, Dynamical evolution of basaltic asteroids outside the Vesta family in the inner main belt, Astronomy and Astrophysics 672 A97 (2023). DOI: 10.1051/0004-6361/202245678.
[59] S. Kashuba, et al., The Simeiz plate collection of the ODESSA astronomical observatory, in: 11th Bulgarian-Serbian Astronomical Conference, 2018, pp. 207-216.

УДК 004.853

КОРИСНИЙ ІНСТРУМЕНТ ТЕЛЕСКОП ДЛЯ ОТРИМАННЯ АСТРОНОМІЧНИХ МЕТАДАНИХ З ФАЙЛОВ FITS

Ph.D. Хламов С.

Харківський національний університет радіоелектроніки, Україна
E-mail: sergii.khlamov@gmail.com

Анотація. *Глава присвячена реалізації підходу інтелектуального аналізу даних щодо метаданих астрономічних файлів із великих архівів. Кожен астрономічний файл має загальноприйнятуструктуру, яка містить спеціальний формат метаданих. Такі метадані містять необхідну астрономічну інформацію, яка необхідна для належного зберігання, аналізу даних, обробки, аналізу під час дослідження. Ця реалізація була реалізована як інструмент під назвою «Телескоп» з використанням мови програмування C#, платформи .NET, технології Windows Forms та оснащена файлом бази даних MDB для СУБД Microsoft Access. Інструмент має два режими: режим консолі для автоматизованої інтеграції з конвеєрами обробки та режим із графічним інтерфейсом користувача (GUI) для візуалізації обробки та додаткових корисних функцій. Інструмент Telescope був розроблений для видобутку великих астрономічних даних з різних архівів, аналізу метаданих з кожного астрономічного файлу та збору їх із подальшим вставленням у базу даних. Такі аналізовані дані використовувалися для різних цілей обробки астрономічних зображень і машинного зору. Інструмент Telescope був розроблений під час досліджень у рамках проекту CoLiTec і протестований з астрономічними файлами з кількох архівів у різних обсерваторіях. Також в таких обсерваторіях було успішно впроваджено та встановлено на конвеєрах обробки астрономічних зображень інструмент Telescope.*

Ключові слова: *інтелектуальний аналіз даних, великі дані, метадані, база даних, потік даних, обробка зображень, машинне бачення, CoLiTec*

UDC 004.853

СИСТЕМА ОЦІНКИ ТА АНАЛІЗУ ТЕКСТОВОГО КОНТЕНТУ НА БАЗІ АЛГОРИТМІВ МАШИННОГО НАВЧАННЯ

Ph.D. Рудніченко М.[a], Ph.D. Шибаєва Н.[a], Ph.D. Отрадська Т.[a], Ph.D. Потієнко О.[a], Ph.D. Шпінарева І. [a], Dr.Sci. Петров І. [b]

[a]Національний університет «Одеська Політехніка», Україна
[b]Національний університет "Одеська Морська Академія", Україна
ORCID: 0000-0002-7343-8076, 0000-0002-7869-9953, 0000-0002-5808-5647, 0000-0002-8740-6198, 0000-0002-0952-2281, 0000-0001-9208-4923
E-mail: nickolay.rud@gmail.com, nati.sh@gmail.com, tv_61@ukr.net, firmn@gmail.com, frumle@ukr.net, iryna.shpinareva@onu.edu.ua

Аннотація. *У роботі визначаються та розглядаються результати розробки та дослідження системи автоматизації процесів оцінки та аналізу текстового контенту на прикладі класифікації текстових статей з урахуванням застосування різних методів машинного навчання. Наведено результати аналізу специфіки та проблематики предметної галузі, особливості обробки природної мови, обґрунтовано актуальність застосування елементів, методів та засобів штучного інтелекту, а також інтелектуального аналізу даних для завдань класифікації тексту. Проведено об'єктно-орієнтоване моделювання програмного забезпечення на базі використання мови UML на прикладі ряду діаграм, описано структуру системи. Розглянуто процедури передобробки та очищення наборів даних для проведення дослідження, виконано чисельні оцінки метрик оцінки точності класифікації 5 різних алгоритмів машинного навчання на навчальній та текстовій вибірках. Проведено процедури підбору гіперпараметрів найбільш підходящої з отриманих моделей з метою підвищення її точності класифікації.*

Keywords: *інтелектуальний аналіз даних, класифікація тексту, аналіз тексту, обробка природної мови, розробка програмного забезпечення, веб-програмування Python*

1. Вступ

У зв'язку з постійним зростанням інформаційних джерел та технічних пристроїв підтримки процесів ознайомлення людини з цікавими даними завдання ефективного аналізу різнорідних текстів є важливим практично в будь-якій сучасній галузі людської діяльності, де його експертиза, досвід та знання можуть бути викладені в текстовому вигляді [1]. З часів винаходу писемності більша частина інформації представлена у вигляді текстів природними мовами. Завдяки розвитку інформаційних технологій створено безліч програмних додатків, бібліотек та допоміжних технічних інструментів, за допомогою яких робляться спроби здійснити комплексну автоматизацію процесів обробки та аналізу текстової інформації [2].

Поява широкого доступу до Інтернету призвела до бурхливому зростанню обсягів доступної текстової інформації, що у відкритому доступі, що значно прискорило розвиток наукових напрямів у галузі обробки тексту, відомих як natural language processing і computational linguistics [3]. У рамках цих напрямків розроблено та запропоновано досить велику кількість перспективних ідей з автоматичної обробки текстів природною мовою, деякі з яких втілені в рамках ряду прикладних систем, у тому числі комерційних. Сфера додатків комп'ютерної лінгвістики постійно розширюється, з'являються нові завдання, які успішно вирішуються, у тому числі із залученням результатів суміжних наукових областей [4]. Найбільш тісно комп'ютерна лінгвістика пов'язана з областю штучного інтелекту, в рамках якої розробляються програмні моделі окремих інтелектуальних функцій з обробки текстових даних [5]. На жаль, потенціал використання подібних засобів і систем недостатньо універсальний, є вузько спеціалізованим під конкретне завдання (розпізнавання тексту, перевірка пунктуації, синтаксису, статистичні підрахунки символів, перевірка унікальності та ін.) більш орієнтований на обробку, ніж на аналіз даних [6]. Не передбачено можливості вибору конкретного методу обробки чи аналізу, дослідження його ефективності, і, головне, внесення модифікацій у його структуру з метою підвищення якості його роботи [7].У зв'язку з цим загальна ефективність і гнучкість є досить низькою, що пов'язано ще й з тим, що повноцінна автоматизація всіх процесів обробки та аналізу даних можлива лише при реалізації логіки розуміння програмним забезпеченням тексту на рівні людини, що вимагає створення повноцінного штучного інтелекту, що не є поки що можливим і здійсненним технічно через брак знань про специфіку процесу сприйняття інформації (зокрема, у текстовому вигляді) людиною [8]. Незважаючи на перетин досліджень у галузі комп'ютерної лінгвістики та штучного інтелекту (що пов'язано з тим, що процес володіння та використання мови для спілкування та передачі інформації відноситься до інтелектуальних функцій людини), їх цільові завдання відрізняються. Узагальнене, ключове завдання комп'ютерної лінгвістики полягає у розробці методів та засобів побудови лінгвістичних процесорів з автоматичної обробки текстів природними мовами. Штучний інтелект більш націлений створення прикладних технічних та інформаційних систем, здатних ефективно і максимально точно вирішувати завдання необчислювального характеру, виконувати дії, що вимагають упорядкованої та осмисленої переробки різнорідної інформації великих обсягів, що є не завжди детермінованою і певною, а також складною для формалізації. В даному ракурсі найбільш доцільним є використання сучасних технологій машинного навчання (МН), за допомогою чого є можливим автоматизація аналізу тексту та вирішення прикладних завдань. Зокрема, перспективним шляхом є застосування контейнеризація та використання конвеєрних засобів розробки моделей штучного інтелекту для аналізу текстів [9]. Конвеєр МН допомагає створювати, оптимізувати та адмініструвати робочий процес створення моделей, він є процесом для прискорення, повторного використання, управління та розгортання моделей МН. Конвеєр МН є наскрізною конструкцією, яка організує потік даних в модель (або набір з декількох моделей) і виведення з неї, включає введення необроблених даних, функції, вихідні дані, моделі та їх параметри, а також вихідні дані прогнозування. У більш широкому сенсі конвеєр МН є способом систематизації та автоматизації робочого процесу, необхідний для створення моделі [10]. Ключова перевага подібних МН конвеєрів полягає в інтегральній автоматизації етапів життевого циклу моделі, зокрема, при оновленні навчальних даних запускаються робочі процеси перевірки, обробки, навчання моделі, тестування та розгортання. Конвеєр МН включає процеси, які [11, 12, 13]:

– ефективно відстежують зміну версій вихідних даних для послідовних запусків нових процесів навчання моделей;

– реалізують попередню обробку даних для навчання та перевірки моделі;

– здійснюють стеження за версіями контрольних точок моделі під час навчання;

– логують експерименти щодо навчання моделей;

– аналізують та перевіряють навчені моделі ML;

– виконують розгортання та масштабування створених моделей;

– оновлюють дані для навчання та реалізують комплексний моніторинг моделі.

Конвеєр МН дозволяє забезпечити наступний ряд переваг: мінімізація помилок; відстеження експериментів; стандартизація всіх етапів; масштабованість моделей на різних архітектурах.

Таким чином, впровадження конвеєрів МН дозволяє досягти наступних результатів: скорочення часу розробки нових моделей, спрощення та передбачуваність процесів оновлення існуючих моделей, скорочення часових витрат на відтворення обчислювальних експериментів. Все це сприяє зниженню витрат на виконання проектів у галузі МН та аналізу даних. Спільним для зазначених наук є використання методів та засобів комп'ютерного моделювання, що є основою проведення досліджень щодо оцінки ефективності застосування окремих підходів для вирішення прикладних практичних завдань, з урахуванням їх евристичних особливостей та характеристик [14]. У зв'язку з цим виникає необхідність розробки та впровадження у процеси аналізу текстових даних гнучких, функціональних та зручних програмних продуктів, що забезпечують як можливості вирішення бізнес-процесів, так і проведення експериментальних аналітичних досліджень, зокрема моніторингу ефективності кожного з імплементованих підходів з можливістю внесення коригувань у його роботу, розширення функціоналу завдяки модульному складу.

В якості конкретної задачі може бути розглянута задача класифікації тексту, яка є однією з найчастіше використовуваних на практиці при автоматизації обробки та сегментації інформаційних матеріалів, що розміщуються на різних веб-ресурсах або електронних документах.

2. Постановка проблеми та аналіз існуючих підходів та публікацій

В даний час, при вирішенні завдань обробки природної мови та класифікації тексту зокрема, для створення модулів лінгвістичних процесорів застосовується два основних підходи: rulebased, в основі якого знаходиться принцип формування правил, що полягає в описі лінгвістичної інформації у вигляді формальних правил і інженерний, заснований МН. У правилах основних систем правила вбудовувалися до складу програмного коду, в даний час для цього використовуються формальні мови, створювані безпосередньо для розробки. Правила створюються лінгвістами або фахівцями з проблемної галузі текстів, що обробляються [15]. У рамках другого підходу, заснованого на МН, джерелом лінгвістичної інформації є відібрані тексти заданої проблемних областей виявлення прихованих закономірностей у принципах побудови пропозицій. На базовому рівні МН можна поділити на два основні типи [16, 17, 18, 19]:

1. МН з учителем (supervised learning) - включає моделювання ознак даних та відповідних даних міток. Після вибору моделі її можна використовувати для присвоєння міток новим невідомим раніше даними. Подібний тип МО поділяється далі на завдання класифікації та регресії. При класифікації мітки є дискретні категорії, а при регресії є безперервними величинами. Алгоритми МН, які навчаються за схемою «об'єкт-відповідь», називають способами навчання з учителем. Це з тим, що вихідне значення відповіді є орієнтиром на формування вихідного значення моделлю МН. Формування набору розмічених даних, тобто. включають об'єкти з відповідними відповідями є вкрай трудомістким процесом, який здійснюється, як правило, найбільш надійним способом, тобто. вручну. При цьому, подібні алгоритми навчання з учителем досить просто інтерпретуються, у зв'язку з чим якість їх функціонування може бути чисельно оцінена.

2. МН без вчителя - включає моделювання ознак набору даних без будь-яких міток за принципом виявлення залежностей автоматично. Такі моделі включають основні завдання: кластеризації (clustering) та зниження розмірності (dimensionality reduction). Алгоритми кластеризації служать виділення окремих груп даних, тоді як алгоритми зниження розмірності призначені пошуку найстисліших уявлень даних. У МН кожен окремий об'єкт чи рядок таблиці називаються прикладом (sample) чи точкою (data point), а стовпці таблиці, формують дані об'єкти називають ознаками (features) чи

характеристиками [20]. Можливі взаємозалежності між даними найчастіше задаються заздалегідь під час обробки даних шляхом формування параметричних вирішальних правил, які оцінюються з урахуванням значень ваг.Конкретні значення ваг, що найчастіше приймають значення від 0 до 1 або -1 до 1, спочатку беруться однаковими або генеруються випадковим чином. Після цього в процесі побудови моделі вони розраховуються при навчанні на основі імпорту навчальної вибірки, яка є деякою кількістю окремих об'єктів з відомими явними та прихованими змінними. Сама задача розрахунку ваг вирішального правила по заданій навчальній вибірці називається завданням налаштування або навчання, а завдання ідентифікації допустимих значень не явною змінною за вказаними явними компонентами об'єкта і вагою - завданням виведення. Найчастіше мається на увазі, кожен окремий об'єкт формується з урахуванням одного набору змінних, причому номенклатура явних і прихованих змінних об'єктів є ідентичною. Переваги сучасних систем з МН у порівнянні з загальноприйнятими альтернативами, такими як ручний аналіз, жорстко задані бізнес-правила та прості статистичні моделі, такі [21, 22, 23]:

– висока точність. МН використовує вхідні дані для створення логіки прикладної програми, призначеної для конкретної задачі. У зв'язку з цим у процесі отримання нових даних підвищується узагальнююча здатність і зростає точність прогнозних значень моделі;

– ефективна автоматизація. У процесі аналізу відповідей та перерахунку ваг модель МН здатна автоматично виявляти у даних кореляцію та нові шаблони. Подібна властивість моделей уможливлює інтеграцію створених систем з МН безпосередньо у вироблені прикладні робочі процеси;

– висока швидкість аналізу. Ефективно побудовані моделі МН дозволяють обробляти та видавати відповіді практично відразу після надходження нових даних шляхом перерахунку вагових значень, що уможливлює використання даного підходу в системах реального часу;

– гнучкість налаштування моделей. У зв'язку з тим, що в основі процесів побудови моделей МН знаходяться дані та математичні алгоритми, аналітики та фахівці в галузі аналізу даних здатні підбирати потрібні параметри для моделей з урахуванням різних набір обмежень, що накладаються предметною областю дослідження;

– масштабованість. У процесі розширення завдань і підвищення обсягів даних логіка моделі МН може бути оперативно адаптована під розподілену архітектуру. Зокрема, ряд алгоритмів МН здатні ефективно обробляти безліч даних на розподілених обчислювальних вузлах у хмарі [24].

Основними та найбільш пріоритетними етапами передобробки даних для використання МН при вирішенні завдань ІАД є [25]: вибірка вхідних даних, полягає у процесі відбору цільових ознак та об'єктів, ґрунтуючись на їх релевантності для конкретних цілей ІАД, у тому числі виходячи з якості та обмежень; очищення даних, полягає у видаленні друкарських помилок у даних, некоректних або відсутніх значень, усунення дублів та різних описів одного об'єкта, відновлення цілісності, унікальності та основних логічних зв'язків; генерація цільових ознак, для їх перетворення на векторні форми для моделей МН, для збільшення рівня точності алгоритмів, що застосовуються; інтеграція, тобто. об'єднання даних із різних типів джерел з їхньою послідовною агрегацією; форматування, що виражається у вигляді синтаксичних змін даних, що не змінюють їх значення та необхідні для забезпечення роботи інструментів моделювання [26].

Застосування методів МН передбачає наявність деякого набору вхідних даних, що застосовуються для навчання класифікатора, і в свою чергу реалізує задані алгоритми. Дані методи перспективніші і мають широкий спектр можливостей у вирішенні завдань класифікації.

Застосування МН особливо ефективно при вирішенні завдань обробки великих колекцій різнопланових документів, що вимагають віднесення кожного текстового фрагмента до певного класу, що може бути зведено до завдань categorization і clustering, які є базовими для направлення text mining. Побудова моделі класифікатор відбувається на спеціально розміченому текстовому корпусі, у якому кожного символу складено мітки, здійснюють функції кодування ознак розпізнавання одиниць текстів. Навчання такої моделі здійснюється за допомогою виявлення загальних закономірностей, властивих текстам природною мовою, на основі dataset [28,37]. Технічно, задача класифікації статей може бути вирішена за допомогою використання існуючих методів класифікації даних на базі використання МН та штучних нейронних мереж. При цьому існує ряд особливостей та суттєвих складнощів, що обмежують та знижують ефективність застосування класичних методів МН для класифікації різноманітних наборів даних [29, 30, 31, 32, 33]. До основних складнощів слід віднести необхідність збирання вихідних текстів з авторитетних джерел, що знаходяться у відкритому доступі, з подальшим очищенням та передобробкою текстових даних для навчання моделі класифікатора, що неможливо зробити вручну, а автоматизація цього процесу може порушити цілісність та семантичний зміст тексту, знижуючи його змістовний потенціал для формування узагальнюючих можливостей моделі [34, 35, 36]. Обидва розглянуті підходи мають свої переваги та недоліки. Процес створення правил є, як правило, ручним, трудомістким і тривалим для лінгвіста, який не завжди може врахувати всі аспекти, винятки та особливості різних мов [37, 38, 39, 40].

Проблематичним є завдання його автоматизації за допомогою використання спеціалізованого прикладного програмного забезпечення, яке не є універсальним і не здатне врахувати різні словосполучення, терміни та словоформи в контексті окремих мовних сімейств та синтаксичних особливостей [41, 42]. Перевагою даного підходу є інтерпретованість та зрозумілість складених правил, що формуються у декларативній формі, що спрощує процес їх підтримки та модифікації [43]. На відміну від правил використовуваного підходу методи МН не вимагають ручної праці зі складання правил, у зв'язку з чим процес обробки даних може бути автоматизований, що сприяє зниженню тимчасових витрат на розробку повноцінних аналітичних систем класифікації [44]. Складності цього підходу полягають у необхідності проведення досить великої кількості обчислювальних експериментів для порівняльного аналізу результатів і виявлення найбільш ефективних методів МН, підбору гіперпараметрів, що впливають на точність віднесення тексту до певного класу.

Додатковими аспектами є складність інтерпретації процесу класифікації та необхідність попередньої розмітки корпусу текстів, що не завжди можливо [45]. У зв'язку з більшим ступенем універсальності і гнучкості для вирішення нами задачі доцільним є вибір МН в якості основи для побудови програмного забезпечення класифікації. Результати аналізу існуючих алгоритмів МН свідчать необхідність створення можливостей для користувача системи побудови різних моделей класифікаторів тексту з метою забезпечення процесу порівняльного аналізу їх ефективності з подальшим підвищенням його точності шляхом інтерактивного підбору гіперпараметрів. Метою роботи є розробка та дослідження системи автоматизації процесів оцінки та аналізу текстового контенту для збільшення ефективності класифікації тексту за категоріями з подальшим виданням рекомендацій для користувачів щодо ознайомлення з контентом за інтересами.

3. Створення проекту системи

З метою визначення та формалізації контексту системи з урахуванням предметної області, що моделюється, доцільно проведення проектування програмного забезпечення. Першим етапом для цього є формулювання загальних вимог до функціональної поведінки за допомогою розробки діаграми ключових прецедентів (варіантів використання), що наведено на рис.1. У рамках проектованої програмної системи користувач зможе здійснювати: завантаження "сирого" набору даних у вигляді набору текстових файлів, що містять у собі текст статті та назву розділу, до якого вона належатиме; перетворення складеного набору даних у структурований вигляд, що має табличний вигляд, більш зручний для подальшого аналізу та роботи; первинний аналіз даних, що включатиме підрахунок та визначення метаданих, у тому числі кількості текстових статей у кожній з категорій, а також розрахунок відсотка текстів за категоріями, розмір статті у форматі txt. Це необхідно для співвідношення даних між собою та формування ознак для подальшого навчання моделей МН; збереження набору даних, що пройшли первинний аналіз у форматі *.csv для його зручної обробки програмним забезпеченням;

очищення текстів, що включає видалення спеціальних символів і розділових знаків, які не потрібні для подальшої класифікації текстів, об'єднання слів, написаних з великої і маленької літери, що несуть однаковий зміст; лематизації та стемінг; розрахунок TF-IDF вектора; створення словників найзначніших слів для класифікації; ініціювання процесу навчання з урахуванням алгоритмів машинного навчання.

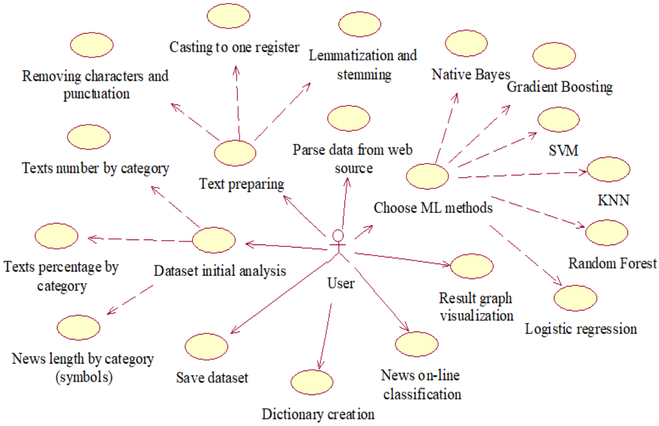


Рис. 1. Діаграма варіантів створеного проекту системи

Вибраними моделями є логістична регресія, наївний класифікатор Байєса, KNN, SVM, Random Forest, Boosting. Навчені моделі зберігаються у форматі *.pickle; вибір кращої моделі створених класифікаторів шляхом порівняння їх точності та ефективності за метриками; подання результатів у вигляді графіків з використанням двох методів зниження розмірності: PCA та t-SNE графік; збирання (парсинг) тексту зі статей, розміщених на вказаних користувачем веб-сайтах, з подальшим їх автоматичним аналізом за допомогою вибраних моделей, що завантажуються в систему динамічно під час операції аналізу; візуалізація результатів класифікації текстів на кожному із веб-сайтів із зазначенням точності роботи кожної моделі.

Більш детальна формалізація динаміки процесу взаємодії користувача із системою наведена на рис.2. Після запуску програмного забезпечення користувач здійснює завантаження даних, які оброблятимуться.

Вибір наборів даних для навчання можливий шляхом підключення різних джерел (формати csv, txt, doc, не захищений pdf, файли баз даних) або з Інтернету (посилання на веб-ресурси або на прямі текстові файли). Усі дані зберігаються у проміжному форматі txt для спрощення процесу подальшої обробки.

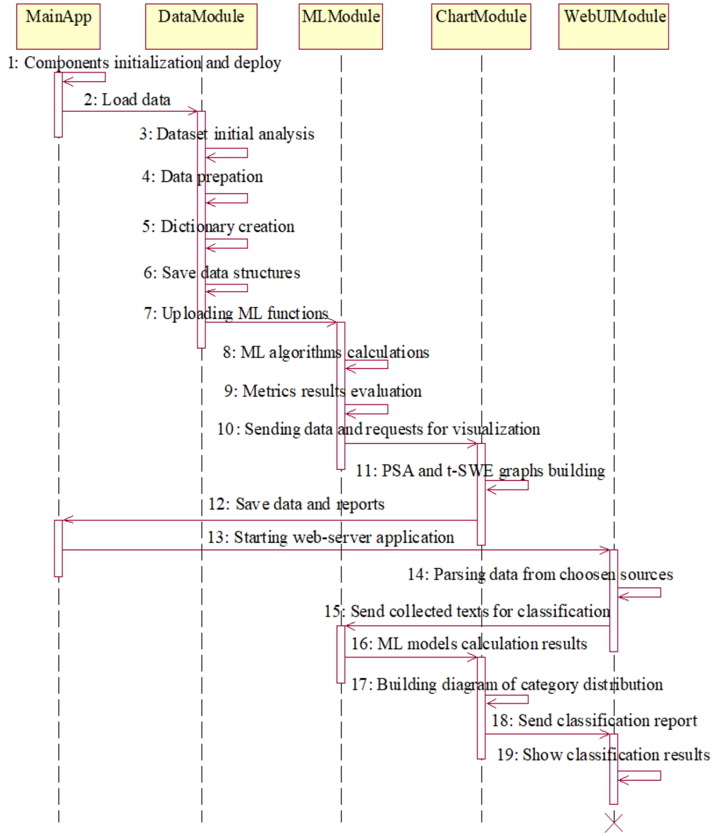


Рис. 2. Діаграма послідовності дій системи

Далі модуль програми, який відповідає за обробку даних отримує dataset, завантажений в систему, формуючи на виході з усіх завантажених наборів тексту зведену таблицю, що містить колонки з назвою файлу, його змістом, категорією (тематикою), до якої відноситься текст. Останній стовпець є вихідний переміреною, може бути розмічений вручну в рамках даного модуля, у разі наявності цього набору в датасеті він може бути позначений за допомогою елементів інтерфейсу. На наступному етапі дані, представлені у вигляді таблиці, проходять первинний аналіз, щоб з'ясувати, наскільки така вибірка є повною для навчання моделей ML і як точно набір даних описує предметну область. Для цього використовується окремий модуль, що базується на аналізі метаданих і детально описаний у рамках роботи [4]. Під час первинного аналізу розраховуються такі показники як підрахунок кількості текстів у кожній із вибраних категорій, відсоток, який складають тексти з кожної рубрики у відношенні до загальної кількості текстів та середній обсяг тексту у кожній із вибраних тематик.

В результаті довгі тексти видаляються з вибірки, що дозволяє збалансувати дані виходячи із вказаних користувачів параметрів та обмежень. Далі проводиться етап очищення даних для початку подальшого аналізу його складу. Саме тоді проводиться зменшення обсягу текстів, виділяються найбільш значимі семантичні елементи. Очищення текстів виконуються шляхом вилучення спеціальних символів і розділових знаків, видаляються слова, що часто повторюються, не несуть істотного змістового змісту (союзи, прислівники) і присвійні займенники, форми однини і множини в однакових словах об'єднуються в одну, проводиться приведення слів до єдиного регістру. Окремо виконуються процеси лематизації та стеммінгу. Наступним етапом є створення пакета словників, що складаються зі слів, які найчастіше трапляються у текстах окремих рубрик, що дозволяє забезпечити семантичну зв'язок між контекстом статей. Оброблені дані та словники зберігаються окремими файлами для подальшого використання у навчанні. Користувач визначає тематики, за якими необхідно провести класифікацію. Далі ініціюються обчислювальні процеси з розбиття вибірок текстових даних на навчальні та тестові, створюються об'єкти класів по кожному з шести використовуваних методів ML. Для проведення обчислювальних процесів у паралельному режимі на кількох обчислювальних вузлах передбачено опцію підключення Apache Mesos та бібліотеки Numba для використання ресурсів графічних апаратних прискорювачів. На основі отриманих результатів обчислень проводиться розрахунок метрик оцінки якості класифікації, головною з яких є Accuracy (ставлення правильних відповідей до загальної кількості примірників текстів). Ще однією метрикою є Precision, як співвідношення правильно позитивних відповідей до всіх позитивних відповідей. Після цього отримані дані оцінки метрик зберігаються у текстовий файл, агрегуються та передаються на вхід графічного модуля, на якому здійснюється візуалізація графіків аналізу точності класифікації. Для більшої міри інтерпретованості результатів необхідним є використання методів зниження розмірності даних. Використані дві методики [1,4,8]: - аналіз основних компонентів, сутність якої полягає у розрахунку власних значень та власних векторів матриці даних для формування мінімальної кількості змінних, які забезпечують максимальне значення дисперсії даних. Використовується у разі коли обсяги тексту не перевищують 10 000 знаків; - t-SNE, у своїй основі містить імовірнісний підхід, що зручно при візуалізації даних великих розмірів (понад 10 000 знаків). Дозволяє мінімізувати розбіжність між розподілом, що вимірює попарно подібність вхідних об'єктів, і розподілом, що вимірює попарно подібність відповідних низькомірних точок вбудовування.

В інтерфейсі користувача розробленого веб-програми містяться поля введення даних та візуалізації у вигляді окремих dashboards. Користувач здійснює вибір веб-сайтів, що цікавлять, з текстовим контентом новин або тематичних статей, які потрібно класифікувати. Автоматичне виконання парсингу заголовків кожної з вибраних новин на вказаній сторінці веб-сайту, а також основного тексту. При розборі кожної сторінки користувач має змогу вибрати якусь частину тексту і які фрагменти він хоче отримати. Для спрощення процесу дослідження відібрано лише англомовні ресурси з лаконічності та зручності даної мови для аналізу. Після аналізу всіх текстових даних за допомогою активації вибраних класифікаторів формується ранжований перелік зі спадання точності для кожного з моделей ML. Після цього відбувається звернення до графічного модуля, з якого виконується побудова двовимірних графіків для аналізу результатів. Описані вище процеси реалізовані у вигляді інтеграції до проекту низки зовнішніх залежностей і створення компонентів. Узагальнена діаграма компонентів розробленого програмного забезпечення наведена на рис.3. Головний модуль програми (MainApp) виконує функції виклику інших модулів програми для обробки запитів програми обробки даних, виконання обчислювальних процесів зі створення моделей ML, побудови графіків і загальної бізнес-логіки роботи програмного забезпечення.

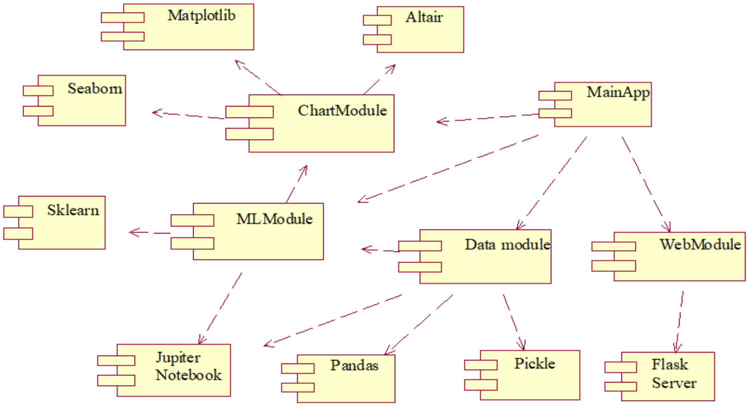


Рис. 3. Діаграма компонентів системи

На підставі створеного проекту програмного забезпечення, що дозволяє позначити його функціонал і призначення доцільним є програмна імплементація засобами мови програмування Python, шляхом наповнення створених класів відповідними методами, полями і властивостями (рис.4).

Для зручності проект системи, що є модульну структуру, у середовищі розробки розділений деякі пакети, у кожному у тому числі містяться класи, реалізують необхідний функціонал з їхньої логічної зв'язності між собою.

Кожен із яких містить свій програмний модуль відповідно.

«Raw dataset» включає необроблені дані у вигляді файлів і логіку реалізації імпорту.

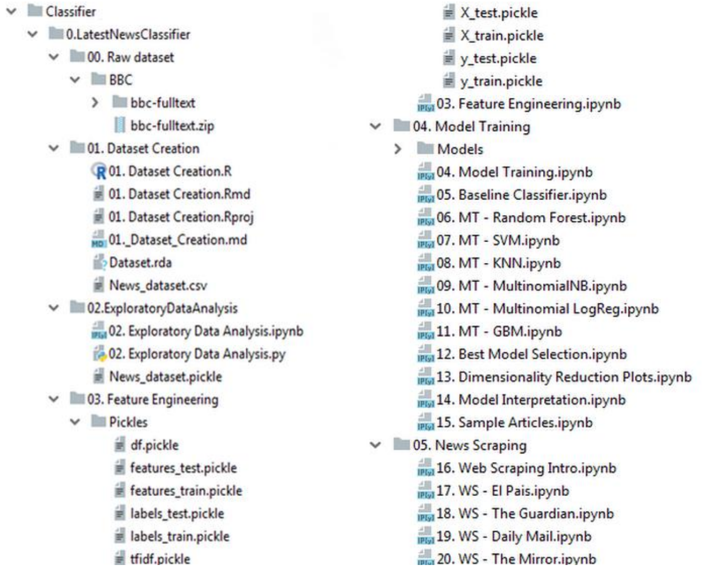


Рис. 4. Структура проекту модулів системи

Dataset Creation забезпечує логіку створення структурованого набору даних. "ExploratoryDataAnalysis" містить класи, що відповідають за первинний аналіз даних. «Feature Engineering» забезпечує очищення даних для подальшої обробки та виділення ознак.Model Training містить відповідні класи, що реалізують виклик і створення об'єктів моделей класифікаторів, а також їх серіалізовані структури, що завантажуються в процесі аналізу даних через веб-інтерфейс. Також у даному пакеті розміщено клас, що містить методи для порівняння всіх класифікаторів та вибору найбільш ефективного для поставленого завдання з урахуванням розрахунку метрик. ScrapingData імплементує логіку збору текстових даних із зазначених сайтів для подальшого аналізу. «UI» реалізує функціонал графічного інтерфейсу користувача та виведення візуалізацій з елементами керування. Для обробки даних обрано бібліотеки: pandas це високорівнева Python бібліотека для аналізу даних. Побудована поверх більш низькорівневой бібліотеки NumPy, що дозволяє підвищити рівень продуктивності. В екосистемі Python, pandas є найбільш просунутою що швидко розвивалася бібліотекою для обробки і аналізу даних. Щоб ефективно працювати з pandas застосовано структури даних бібліотеки: DataFrame і Series. Series вдає із себе об'єкт, схожий на одновимірний масив, але відмітною його рисою в нашому завданні є наявність асоційованих міток, індексів, уздовж кожного елемента зі списку. Така особливість перетворює його в асоціативний масив або словник. Для управління даними побудованих моделей застосовано модуль pickle, що реалізує алгоритми серіалізації і десеріалізациі об'єктів "Pickling" - процес перетворення об'єкта Python в потік байтів, а "unpickling" - зворотна операція, в результаті якої потік байтів перетворюється назад в Python-об'єкт. Бібліотекою для машинного навчання було обрано scikit-learn, що орієнтована в першу чергу на моделювання даних, а не на завантаження, маніпуляцію і узагальнення даних. Зо допомоги бібліотеки нами здійснюється: перехресна перевірка (Cross Validation) для оцінки ефективності роботи моделей на незалежних даних та для тестових наборів даних і генерації наборів даних з певними властивостями при дослідженні поведінкових властивостей моделі, визначення атрибутів в текстових даних.

Для візуалізації розподілу метричних змінних засобами бібліотеки Matplotlib використовуються наступні типи графіків функцій: distplot; jointplot; rugplot; kdeplot.

Візуалізація відносного розподілу між парами змінних виконується за допомогою методів: pairGrid, pairplot, facetGrid. Для парсингу веб сторінок та збору текстів статей було обрано бібліотеку BeautifulSoup з бібліотекою Request для відображання вмісту URL-адреси. Прототипи графічного дашборду виводу даних наведено на рис.5. Система побудована так, що збирає останні новини з головних сторінок, обраних трьох сайтів: «The Guardian», «Sky News», «El Pais England».

У зв’язку з тим, що кожний з цих сайтів має свою власну html структуру сторінки, довелося створювати три окремих модулі(парсери) для кожного з них окремо. Розглянемо роботу кожного з них. На прикладі «El Pais England» спочатку відправляється запит на отримання змісту сайту за допомогою надісланого URL. Нам у відповідь повернеться розмітка сторінки, яку ми зберігаємо, як текст.

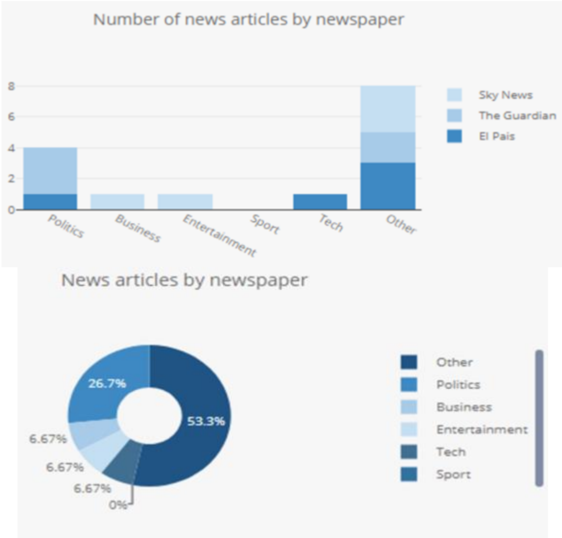


Рис. 5. Прототипи графічного дашборду інтерфейсу системи

Проаналізувавши структуру сайту ми бачимо, що заголовок кожної новини виступає тегом h2 й належить класу 'articulo-titulo', наявність окремого класу для новин значно полегшить нам задачу збору усіх новин.

Запускаючи пошук ми підрахуємо кількість новин, які зараз розміщені на головній сторінці. Їх 69 згідно заголовків. Далі формується список, який складається із заголовків новин і посилань на самі тексти. Використовуючи утворений список дістаємо тексти новин та формуємо два набори даних: набір даних, який буде вводитися до моделі (df_features); набір даних із заголовком та посиланням (df_show_info). Метрикою для модуля з парсингом сайту є час, який витрачається на збір новин та їх даних. Для цього усі дії над сайтом та новинами помістили в одну функцію, яка потім буде викликана. Отриманим результатом є значення 2.640945 секунд, що вважається добрим показником. Мінімізувати час важливо для зручності користувачів у майбутньому. Для сайту «The Guardian» виконали усі ті ж пункти. Проаналізувавши виявили, що назви новин знаходяться у тезі h3 й належить класу 'fc-item__title'. У результаті було отримано 95 текстів та посилання на них. У результаті час затрачений на збір даних з сайту «The Guardian» виявився меншим і дорівнює 1.905263секунд. Для сайту «Sky News» було повторено шаги попередніх новинних порталів, назви новин знаходяться у тезі h3 й належить класу 'linkro-darkred'. У результаті було отримано 181 текст та посилання на них. Так само сформовано два набори даних, які об’єднані по таблицях. Час витрачений на цей сайт дорівнює 11.067029, що значно довше за попередні результати, але все ж таки допустимо. Таку затримку по часу можна пояснити складним інтерфейсом, що грузиться довше, та більшою кількістю новин

4. Дослідження роботи системи

У якості вихідних даних взяті набори статей із веб-сайтів qz, theguardian, Feedly. Етапами дослідження стали збір і обробка даних, відбір статей за найбільш популярними рубриками для навчання моделей МН і оцінка точності класифікації з підбором найбільш підходящих значень гіперпараметрів для кращого з виявлених класифікаторів. На базі аналізу зібраної метаінформації по статтям було підібрано п'ять основних категорій статей: бізнес, розваги, політика, IT і здоров'я. За результатами підрахунку кількості текстів у кожній із категорій були побудовані кожну графіку (рис.6), що дозволяє наглядніше відобразити дані (відсоток із категорії у виборі).

Відповідно до графіка можна бачити, що більше текстів в категоріях IT, і менше всього в залученнях, але незважаючи на те, що ці дані залишаються збалансованими, що дозволяє уникнути додаткових маніпуляцій при обробці даних. Далі були проведені процедури обробки та очищення тексту, створення словників, а також ряд обчислювальних процесів. Зокрема, виділено наступні спеціальні символи, що містяться в тексті: \r, \n, 's (апостроф у притягальних місцях (наприклад, government's = уряд \ 's), ' (апостроф у притягальних місцях в інших граматичних формах, " (цитати). Усі ці символи були виділені з тексту та замінені на символ символу з даних. У процесі проведення дослідження вибрано випадкове розбиття набору даних: 85% спостережень, складових тренувальний набір і 15% спостережень, складових тестовий набір. Процес налаштування гіперпараметрів моделі був виконаний з перехресною валідацією тренувального набору, щоб кожна модель класифікатора стала більш узагальнюючою.

Наступним етапом був проведений розрахунок метрики TF-IDF для визначення того, наскільки важливо кожне слово в тексті.

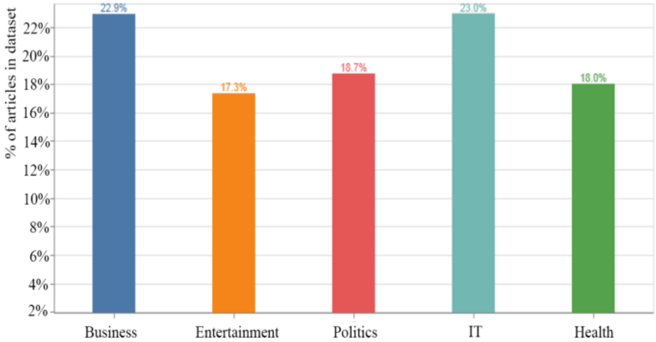


Рис.6. Діаграма частки текстів у кожній із визначених категорій

Процес оформлення полягає в наступному: якщо слово зустрічається в будь-якому документі часто, при цьому зустрічається рідко у всіх інших документах, то це слово має найбільшу значимість для цього документа. Далі відзначено, що всього виділено близько 300 слів для кожної теми з теми. Виконується шляхом виклику методу TfidfVectorizer(). Від початку обробляється лише тренувальний набір даних, тестовий при цьому залишається незмінним. В результаті ми отримаємо слова та словосполучення, які виступають найбільш значущими для кожної з категорій. Формування словників з окремих слів для кожної тематики статей дозволило забезпечити найкращу кожну (чим при звичайних словосполученнях) відповідність окремих слів семантичному значенню в рамках категорії. Після навчання класифікаторів і оцінки роботи моделей з існуючими категоріями була додана додаткова категорія (інше), у разі якщо стаття має низьку належність до існуючого типу (рис.7). Встановлено, що якщо при подачі на вхід до системи тексту невизначеної категорії (семантичні не належать розглянутій п’яти), то класифікатор ставив її до невірної з точністю близько 89%. Наступною величиною, яку визначено в процесі аналізу даних є довжина тексту кожного окремого запису. Довжиною тексту визначено кількість символів, що міститься в кожній із новинних статей. Було розраховано, що лише 7 текстів містить більше 10 000 символів. Щоб виразити залежність більш повно вилучено найдовші тексти з вибірки (рис.8).

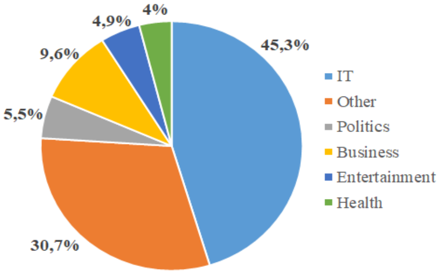


Рис. 7. Відсоток текстів для кожної категорії

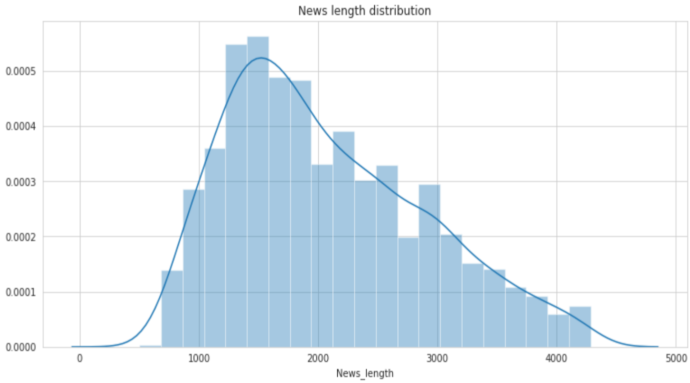


Рис. 8. Розподіл довжини текстів зі зміненим масштабом

На основі тих самих даних було побудовано діаграму розмаху Box Plot - це зручний спосіб візуального представлення груп числових даних через квартилі для нашого датасету (рис.9). Встановлено, що хоча розподіл довжини текстових записів різний для кожної категорії, різниця не надто велика та е вносить додаткового шуму у дані.

Якби у нас були занадто різні довжини між категоріями, у нас виникла б проблема, оскільки процес створення функції може враховувати кількість слів. Однак, створюючи функції за допомогою

TF-IDF оцінювання, ми нормалізуємо функції лише для того, щоб уникнути цього. У процесі моделювання були підібрані гіперпараметри для підвищення якості моделі класифікатора, які показали найкращі результати. Була складена таблиця можливих значень і проведений рандомізований пошук, використовуючи 3-кратну перекрестну перевірку (з 50 ітераціями). Один цикл кросвалідації включає розбиття набору даних на частини, потім побудова моделі на одній частині (званої тренувальним набором), і валідація моделі на іншій частині (званої тестовим набором).

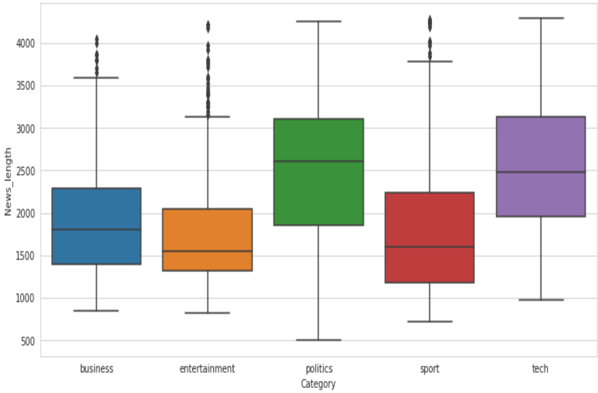


Рис. 9. Діаграма розмаху довжини текстів по категоріям

Щоб зменшити розкид результатів, різні цикли кросвалідації проводяться на різних множинах, а результати валідації усереднюються по всіх циклах. Для цього використано метод бібліотеки sklearn train_test_split(). У якості параметрів функції було передано контент тексту, мітку категорії, до якої належить текст, розмір тестової вибірки (тобто у нашому випадку 15%) і показник random_state, який визначить випадковою величиною індекси наших тренувальних та тесту вальних вибірок. У результаті метод поверне нам чотири масиви x_train, y_train, x_test, y_test. Після отримання моделей з кращими гіперпараметрами, проводився пошук по таблиці на базі за допомогою 3-кратної перехресної перевірки в центрі цих значень, визначати найкращу комбінацію з усіх. Перевагою даного підходу є те, що за допомогою обмеженого рандомізованого пошуку ми можемо охопити набагато більш широкий діапазон значень для кожного гіперпараметра, що істотно не підвищує затрати часу на визначальні експерименти. У якості метрики оцінки моделей класифікаторів використовуються показники точності, точності, запам'ятовування (використовується для вимірювання долі позитивних зразків, правильно класифікованих) і ключ F1-Score (представляє собі середнє значення гармонії між значеннями запам'ятовування та точності). Для випадкового лісу отримуємо наступні значення гіперпараметрів{'bootstrap': False, "max_depth" 40, "max_features": "sqrt", "min_samples_leaf ': 1,' min_samples_split '5,' n_estimators "800}. Саме в цих значеннях модель демонструє найкращі результати. Середня точність моделі з цими гіперпараметрами: 0,845. Точність на тренувальному наборі даних рівна 1.0. Точність на тестовому наборі даних - 0,83. Наступним було розглянуто метод k-ближніх сусідів, ключовим параметром якого є `K` кількість сусідів, які використовуються за умовчанням для запитів. Найбільш високе значення показника точності було отримано при K = 6, середня точність = 0,829. Значення точності на тренувальному наборі даних досягло 0,839, а на тесті трохи менше, що становить 0,818. Для моделі наївного байсовського класифікатора гіперпараметри використовувалися за замовчуванням: alpha = 1.0, class_prior = None, fit_prior = True. На навчальній множині даних був досягнутий результат accuracy = 0,813, тоді як на тестовому виборі accuracy = 0,784. Після того, як були отримані результати класифікації для всіх створених моделей МН, було виконано їх порівняння та обрано найбільш точний. Результати наведені в таблиці 1. Дані відсортовано за зменшенням точності на тестовому наборі даних. Після дослідження при навчанні класифікатору на текстах з визначеними категоріями та подачі різних статей, які не належать до жодної з 5 категорій, ми встановили, що якщо траплявся текст із невизначеної раніше категорії, то класифікатор відносив її до невірної з точністю не перетнувши поріг у 65%. Згідно таблиці 1 видно, що Gradient Boosting Machine є перевченим. Наївний байєсовский класифікатор і метод найближчих сусідів показують менш значні результати, проте в цілому досить ефективні.

Таблиця 1

Точність роботи створених моделей МН

	Model name	Train dataset accuracy	Test dataset accuracy
1	Gradient Boosting	1.0	0.878
2	SVM	0.929	0.919
3	Naive Bayesian	0.813	0.784
4	KNN	0.839	0.818
5	Random Forest	0.845	0.83

Проблема перенавчання є рідкістю в МН, тобто. модель показує хороший результат на тренувальному наборі даних і значно вищий на тестовому.

У зв'язку з цією перевченою моделлю можна вилучити для забезпечення більшої точності або провести процедуру регулярізації. Найбільш точні результати класифікації методу опорних векторів, оскільки мають саму високу точність тестового набору та дуже низький розкид між тестовим і тренувальним вибірками.

Далі проведена підбірка гіперпараметрів для даного класифікатора з метою вивчення потенціалу підвищення його точності. Існують різні типи моделей SVM, ми розглянули лінійні, сигмовидні та RBF. Результати оцінки різних типів моделей SVM залежно від параметра C у таблиці 2.

Проведено 13 окремих обчислювальних експериментів. В процесі виконання досліджень було зафіксовано перенавчання ряду моделей, що часто зустрічається коли модель показує добрий результат на тренувальному наборі даних і відносно гірший на тестовому. У зв’язку з цим ми вилучали з вибірки перенавчені моделі, а найточніший результат має метод опорних векторів, оскільки має найвищу Accuracy тестового набору.

Визначено accuracy при порівнянні моделей та при виборі найкращих гіперпараметрів. У першому випадку ми обчислили accuracy як на навчальних, так і на тестових наборах, щоб виявити моделі надлишків.

Однак ми також отримали матрицю помилок (confusion matrix) та звіт про класифікацію (який обчислює precision, recall та F1-Score для всіх класів) для кожної моделі, щоб ми могли далі інтерпретувати їх поведінку.

Таблиця 2

Значення різних типів моделей SVM

		param_kernel = linear		param_kernel = rbf		param_kernel = sigmoid	
№	param_C	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score
0	0,1	0,949	0,901	0,961	0,896	0,882	0,784
1	0,2	0,963	0,924	0,969	0,927	0,903	0,848
2	0,3	0,968	0,948	0,968	0,918	0,915	0,844
3	0,4	0,965	0,945	0,969	0,928	0,922	0,887
4	0,5	0,968	0,929	0,969	0,947	0,893	0,842
5	1	0,965	0,936	0,969	0,936	0,912	0,831
6	1,5	0,961	0,942	0,97	0,962	0,889	0,799
7	2	0,959	0,93	0,963	0,908	0,913	0,841
8	2,5	0,958	0,929	0,956	0,921	0,91	0,855
9	3	0,955	0,926	0,939	0,887	0,875	0,831
10	3,5	0,953	0,953	0,9	0,834	0,856	0,806
11	4	0,952	0,949	0,823	0,759	0,831	0,791
12	4,5	0,952	0,941	0,745	0,653	0,813	0,751
13	5	0,95	0,932	0,71	0,614	0,806	0,749

Найкращий результат точності (0,97) отримано в моделі RBF при C = 1,5. Графік залежності точності від співвідношення гамма наведено на рис.10. Графік показників F1 для кожного param_kernel із C наведено на рис.11.

У таблиці 3 наведено результати додаткових досліджень моделей для 10 сценаріїв наборів параметрів. RBF модель методу опорних векторів показала кращий результат за лінійну.

Очевидно, що використання досліджених параметрів в певній комбінації дозволяє підвищити рівень якості класифікації текстових новин.

Порівнявши різні класифікатори найкращий результат з найвищою точністю показав метод опорних векторів. Значення Accuracy досягло 0.940120, але підібравши точніше коефіцієнти нам вдалося підвищити значення цієї метрики до 0.970133333, що в порівнянні з іншими класифікаторами є вищим показником, що свідчить про можливості тюнінгу моделей у подальшому.

Таблиця 3

Значення додаткових досліджень за 2 параметрами

	param_gamma	param_kernel	ACCURACY
1	0,1	rbf	0,949
2	0,2	rbf	0,963
3	0,3	rbf	0,968
4	0,4	linear	0,965
5	0,5	linear	0,968
6	1	linear	0,965
7	1,5	sigmoid	0,961
8	2	sigmoid	0,959
9	2,5	sigmoid	0,958
10	3	sigmoid	0,955

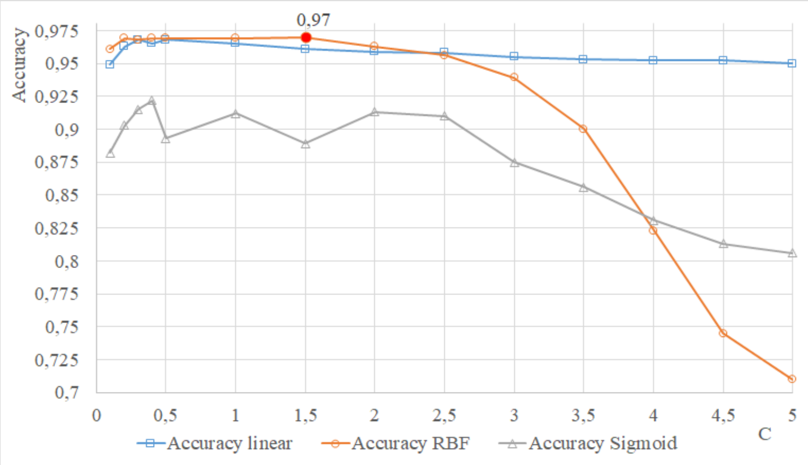


Рис. 10. Графік точності для кожного param_kernel з С

Слід зазначити, що найвищий результат метрики досягається за допомогою RBF, точність = 97% і F1 Score = 96,2% зі значенням C = 1,5. Близькі, але трохи менші значення (точність близько 96,5% і оцінка F1 близько 94,3) були отримані на основі використання лінійної функції, найгірші результати демонструє сигмоїдна функція (точність близько 92,6% і оцінка F1 близько 89,4), яка вказує на наявність хоч і не надто великого, але існуючого зв’язку між значеннями цих параметрів та їх впливом на кінцеву точність моделі.

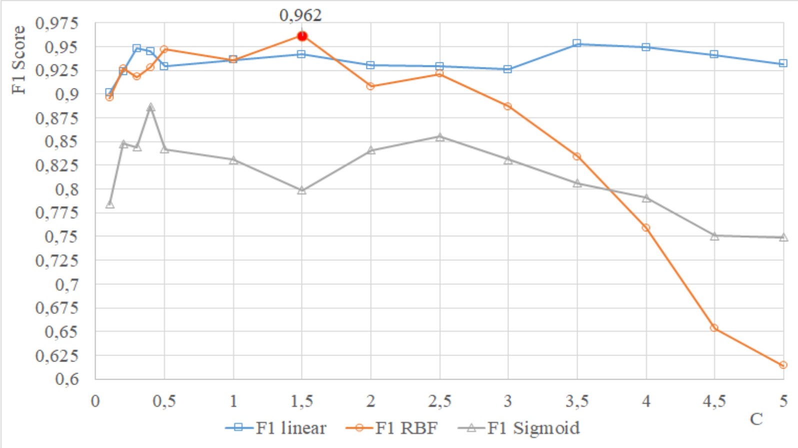


Рис. 11. Графік показників F1 для кожного param_kernel із C

5. Висновки

В результаті даного дослідження було розроблено програмне забезпечення автоматизації класифікації текстових станів різної тематики на базі використання алгоритмів машинного навчання та проведено його тестування. Встановлено, модель SVM на базі функцій RBF є найкращим результатом за лінійну. Це свідчить про те, що задача пошуку можливих комбінацій гіперпараметрів є актуальною, т.к. правильно підібрані значення дозволяють підвищити рівень якості класифікації текстових станів. Завдяки описаному підходу точність класифікації була збільшена в середньому на 5-6%. У подальшому даний процес може бути автоматизований за рахунок впровадження можливостей ітеративного прогону моделей у вибраному діапазоні у багатопоточному режимі, що може забезпечити суттєве зниження тимчасових витрат на проведення вимірних процесів.

6. Література

[1] N. Rudnichenko, V. Vychuzhanin, N. Shybaieva, D. Shybaiev, T. Otradskaya, I. Petrov, The use of machine learning methods to automate the classification of text data arrays large amounts. Information management systems and technologies. Problems and solutions. Ecology, Odessa, (2019) 31-46.

[2] V.V. Vychuzhanin, N.D. Rudnichenko, Metody informatsionnykh tekhnologiy v diagnostike sostoyaniya slozhnykh tekhnicheskikh sistem. Monografiya, Odessa (2019).

[3] N. Rudnichenko, S. Antoshchuk, V. Vychuzhanin, A. Ben, I. Petrov, Information System for the Intellectual Assessment Customers Text Reviews Tonality Based on Artificial Neural

Networks. Proceedings of the 9th International Conference "Information Control Systems & Technologies" (2020) 371-385.
[4] N. Rudnichenko, V. Vychuzhanin, I. Petrov, D. Shibaev, Decision Support System for the Machine Learning Methods Selection in Big Data Mining. Proceedings 0f The Third International Workshop on Computer Modeling and Intelligent Systems (CMIS-2020): session 6 "Intelligent Information Technologies (2020) 872-885.
[5] N. Rudnichenko, S. Antoshchuk, V. Vychuzhanin, A. Ben, I. Petrov, Information System for the Intellectual Assessment Customers Text Reviews Tonality Based on Artificial Neural Networks. Proceedings of the 9th International Conference "Information Control Systems & Technologies" (2020) 371-385.
[6] N.D. Rudnichenko, V.V. Vychuzhanin, N.O. Shibaeva, D.S. Shibaev, T.V. Otradskaya, I.M. Petrov, Software development for interactive big data mining based on machine learning methods. Actual problems of information systems and technologies: monograph (2020) 59-71.
[7] W. Lidong, A, Cheryl, Machine Learning in Big Data. International Journal of Mathematical, Engineering and Management Sciences 1 (2016) 52-61. doi:10.33889/IJMEMS.2016.1.2-006.
[8] Q. Junfei, Wu. Qihui, D. Guoru, Xu. Yuhua, F. Shuo, A survey of machine learning for big data processing. EURASIP Journal on Advances in Signal Processing (2016). doi: 10.1186/s13634-016-0355-x.
[9] B.O. Bliznyuk, L.V. Vasilieva, I.D. Strelnikov, D.S. Tkachuk, Modern methods of natural language processing. Visnik of Kharkiv National University of the Name of V.N. Karazin (2017).
[10] V.S. Yuskov, I.V. Barannikova, Comparative Analysis of Natural Language Processing Platforms. Mining information and analytical bulletin 3 (2017) 272-278.
[11] N.I. Widiastuti, Deep Learning – Now and Next in Text Mining and Natural Language Processing. IOP Conference Series: Materials Science and Engineering 407 (2018). doi:10.1088/1757-899X/407/1/012114.
[12] H.Li, Deep learning for natural language processing: Advantages and challenges. National Science Review 5 (2018). doi:10.1093/nsr/nwx110.
[13] D.V. Noskov, Text classification using machine learning algorithms. Bulletin of Science and Education. №40 (2018) 39-41.
[14] T.V. Batura, Methods for automatic classification of texts. Software products and systems. №30 (2017) 85-89. doi: 10.15827 / 0236-235X.030.1.085-099
[15] W. Yang, J. Boyd-Graber, P. Resnik, A discriminative topic model using document network structure. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (2016) 686–696.
[16] D.A. Adeniyi, Z. Wei, Y. Yongquan, Automated web usage data mining and recommendation system using K-Nearest Neighbor (KNN) classification method. Applied Computing and Informatics (2016) 90–108.
[17] J. Lilleberg, Y. Zhu, Y. Zhang, Support vector machines and word2vec for text classification with semantic features. Cognitive Informatics & Cognitive Computing 2015 IEEE 14th International Conference (2015) 136–140.
[18] A. Jain, J. Mandowara, Text classification by combining text classifiers to improve the efficiency of classification. International Journal of Computer Application. №2 (2016).
[19] M.N. Krasnyansky, A.D. Obukhov, A.A. Voyakina, E.M. Solomatina, Comparative analysis of machine learning methods for solving the problem of classifying documents of a scientific and educational institution. Voronezh State University Bulletin. Series: Systems Analysis and Information Technology. №3 (2018) 173-182 10.17308/sait.2018.3/1245.

[20]O. Abikoye, S. Omokanye, T. Aro, Text Classification Using Data Mining Techniques: A Review. Information Systems Education Journal. №22 (2018) 1-8.
[21]J. Han, H. Ji, Y. Sun, Successful Data Mining Methods for NLP. (2015) 10.3115/v1/P15-5001.
[22]U. Sivarajah, M. Mustafa Kamal, Z. Irani, V. Weerakkody, Critical analysis of Big Data challenges and analytical methods. Journal of Business Research 70 (2017) 263-286.
[23]P. Bala, Introduction of Big Data With Analytics of Big Data. Advanced Deep Learning Applications in Big Data Analytics (2021) 110-125.
[24]J. Won, K. Gun, L. Jiwon, L. Hyunwuk, H. Dongho, R. Won, Deep learning with GPUs. Advances in Computers (2021) doi: 10.1016/bs.adcom.2020.11.003.
[25]H. Ryan, K. Panagiotis, P. Charalampos, T. Lazaros, L. Xing, L. George, A Prototype Deep Learning Paraphrase Identification Service for Discovering Information Cascades in Social Networks. IEEE International Conference on Multimedia and ExpoAt (2020). doi: 10.1109/ICMEW46912.2020.9106044.
[26]S. Will, Neural Networks in Big Data and Web Search. NEURAL NETWORKS IN WEB SEARCH 4 7 (2018) doi:10.3390/data4010007.
[27]L. Zichao, J. Xin, S. Lifeng, L. Hang, Paraphrase Generation with Deep Reinforcement Learning. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (2018) 3865-3878. doi:10.18653/v1/D18-1421.
[28]W. Qiang, L. Bei, X. Tong, Z. Jingbo, L. Changliang, W. Derek, C. Lidia, Learning Deep Transformer Models for Machine Translation (2019) 1810-1822. doi:10.18653/v1/P19-1176.
[29]V.S. Yuskov, I. V. Barannikova, Comparative Analysis of Natural Language Processing Platforms. Mining information and analytical bulletin 3 (2017) 272-278.
[30]H. Hassan, C. Beneki, S. Unger, M. Mazinani, M. Yeganegi, Text Mining in Big Data Analytics. Big Data and Cognitive Computing 4 1 (2020). doi:10.3390/bdcc4010001.
[31]S. Zhou, Research on the Application of Deep Learning in Text Generation. Journal of Physics: Conference Series 1693 (2020). doi:10.1088/1742-6596/1693/1/012060.
[32]A.O. Soshnikov, The role of the functional-semantic type of speech in the process of automatic processing of natural language. Eurasian Union of Scientists 15 (2015) 134-135.
[33]C. Hou, C. Zhou, K. Zhou, J. Sun, S. Xuanyuanj, A Survey of Deep Learning Applied to Story Generation (2019). doi:10.1007/978-3-030-34139-8_1.
[34]T. Young, D. Hazarika,S. Poria, E. Cambria, Recent Trends in Deep Learning Based Natural Language Processing. IEEE Computational Intelligence Magazine 13 (2018) 55-75.
[35]A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. Gomez, Ł. Kaiser, I. Polosukhin, Attention is all you need. Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17). NY, USA (2017).
[36]Y. Wu, L. Liu, Z. Xie, K. Chow, W. Wei, Boosting Ensemble Accuracy by Revisiting Ensemble Diversity Metrics, in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp.16469–16477.
[37]F. Rashidi, S. Nejatian, H. Parvin, V. Rezaie, Diversity based cluster weighting in cluster ensemble: an information theory approach, Artificial Intelligence Review, 52 (2019), 1341–1368. DOI:10.1007/s10462-019-09701-y.
[38]R. Burton, S. Cuff, M. Morgan, A. Matt, M. Eberl, GeoWaVe: Geometric median clustering with weighted voting for ensemble clustering of cytometry data, Bioinformatics (Oxford, England), (2022). DOI:39. 10.1093/bioinformatics/btac751.
[39]Z. Wang, H. Parvin, S. Qasem, B. Tuan, K. Pho, Cluster ensemble selection using balanced normalized mutual information, Journal of Intelligent & Fuzzy Systems, 12 (2020) 1–23. DOI: DOI:10.3233/JIFS-191531

[40] S. Bi, W. Liu, Clustering Analysis of Online Teaching Cases and Evaluation of Teaching Results, International Journal of Emerging Technologies in Learning (iJET), 18 (2023) 128-142. DOI:10.3991/ijet.v18i03.38055.
[41] H.Li, X. Ye, A. Imakura, T. Sakurai, LSEC: Large-scale spectral ensemble clustering. Intelligent Data Analysis, 27 (2023) 59-77. DOI:10.3233/IDA-216240.
[42] T. Uçar, A. Karahoca, Benchmarking data mining approaches for traveler segmentation, International Journal of Electrical and Computer Engineering (IJECE), 11 (2021) 409-416. DOI: 10.11591/ijece.v11i1.pp409-415.
[43] S, Rogic, L. Kascelan, Segmentation Approach for Athleisure and Performance Sport Retailers Based on Data Mining Techniques, International Journal of E-Services and Mobile Applications, 13 (2021) 71-85. DOI:10.4018/IJESMA.2021070104.
[44] A. Gunandi, H. Awang, E. Alhawad, L. Shabaan, Customer Value and Data Mining in Segmentation Analysis, International Journal of Information Technology and Computer Science Applications, 1 (2023) 20-34. DOI:10.58776/ijitcsa.v1i1.16.
[45] C. González, R. Delgado, J. María, S. Santos, Segmentation of Potential Fraud Taxpayers and Characterization in Personal Income Tax Using Data Mining Techniques, Revista Hacienda Pública Española, 239 (2021) 127-157. DOI:10.7866/HPE-RPE.21.4.4.

UDC 004.853

SYSTEM FOR EVALUATION AND ANALYSIS OF TEXT CONTENT BASED ON MACHINE-BASED ALGORITHMS

Ph.D. Rudnichenko M.[a], Ph.D. Shibaeva N.[a], Ph.D. Otradska T.[a], Ph.D. Potienko O.[a], Ph.D. Shpinareva I.[a], Dr. Sci. Petrov I.[b]

[a]National University “Odeska Polytechnic”, Ukraine
[b]National University "Odeska Maritime Academy", Ukraine
E-mail: nickolay.rud@gmail.com, nati.sh@gmail.com, tv_61@ukr.net, firmn@gmail.com, frumle@ukr.net, iryna.shpinareva@onu.edu.ua

Abstract. *The work shows and examines the results of developing and researching a system for automating evaluation processes and analyzing text content in the application of classifying text articles in the context of various machine learning methods. The results of the analysis of the specifics and problems of subject matter, the peculiarities of processing natural language, the relevance of the consolidation of elements, methods and features of individual intelligence, as well as the intelligent analysis of data for the assignment of classification are determined. An object-oriented modeling of software based on the UML language has been carried out using a number of diagrams, and the structure of the system has been described. The procedures for reprocessing and purification of data sets for further investigation are reviewed, numerical estimates of metrics for assessing classification accuracy of 5 different machine-based algorithms on initial and text selections are determined. A procedure was carried out to select the hyperparameters most suitable from the selected models by increasing the accuracy of classification.*

Keywords: *intelligent data analysis, text classification, text analysis, natural language processing, software development, Python web programming,*

UDC 004.853

HELPFUL THRESHOLD TOOL FOR DATA MINING OF ASTRONOMICAL PROCESSING CONFIGURATION PARAMETERS

Ph.D. Khlamov S.

Kharkiv National University of Radio Electronics, Ukraine
ORCID: 0000-0001-9434-1081
E-mail:: sergii.khlamov@gmail.com

Abstract. *The chapter is devoted to implementation of the data mining purposes related to the astronomical processing configuration parameters during invocation of the image processing pipeline. Each processing module in the pipeline has a lot of configuration parameters, which allow tuning the processing process as well to improve the accuracy and speed of calculations during processing. Because of the big amount of such configuration parameters in the complex pipeline of scientific software, the data mining approach is very useful and productive. For our research we have selected scientific software for detection of the moving objects in a series of CCD-frames called "CoLiTec". Such software performs a lot of different image processing tasks, like filtering, background alignment, objects detection, astrometry, photometry, motion detection, etc. The CoLiTec software consists of more than 30 mathematical and processing modules related to the different stages of image processing, where each of them has a lot of configuration parameters to be set. So, to resolve the management issues of astronomical processing configuration parameters, the developers decided to create a new tool for the data mining of such parameters called "ThresHolds". It was implemented as software with graphical user interface (GUI) using the Java programming language, JavaFX technology. The main goal of ThresHolds tool is to classify configuration parameters, manage and validate them, visualize for the end user, and prepare the batch of required parameters for the appropriate stage in processing pipeline. The ThresHolds tool in scope of the CoLiTec software was successfully installed as the main astronomical image processing pipeline in the different observatories.*

Keywords: *Data mining, big data, processing pipeline, dataflow, configuration parameters, image processing, machine vision, CoLiTec*

1. Introduction

There are different tasks for the astronomical image processing and machine vision purposes in astronomy. Some of them are as follows: filtering [1], brightness equalization [2], background alignment [3], object and motion detection [4, 5], astrometry, photometry, images cross-matching, object recognition [6], Wavelet coherence analysis [7] and others. Such astrophysical objects that can be processed and detected in the series of CCD-images [8] are as follows: galaxies, stars, robots [9, 10], drones [11], rockets, satellites [12], and even comets or asteroids [13]. The complex scientific processing pipelines are required to implement the different image processing and machine vision tasks. In common words, the processing pipeline is a set of data processing modules connected in series, where the output of one module is the input of the next one [14]. The modules in pipeline are often executed consequently one by one and rare in parallel or in time-sliced fashion. Also, there are different buffer storage can be inserted between the modules to save the intermediate results or processing data. There are different processing pipelines in the software engineering and computer science [15]:

– instruction pipeline is a classic pipeline that is used in the microprocessors and central processing units (CPUs) to allow overlapping execution of multiple instructions with the same circuitry. Such circuitry is often divided up into stages and each stage is responsible for the processing of a specific part of one instruction at a time and passing results to the next stage (for example, instruction for the coding/decoding/encoding, logic/arithmetic or register fetch).

– graphics pipeline is a pipeline in the most graphics processing units (GPUs), which contains the multiple arithmetic units for implementing the different stages of rendering operations (for example, window clipping, color and light calculation, perspective projection, rendering, etc.).

– software pipeline is a sequence of computing processes (program runs, commands, tasks, procedures, threads, etc.) that executed in parallel, with the output stream of one process being automatically fed as the input stream of the next one.

Along with the software processing pipeline the data pipeline, known as a dataflow, is also performed. And when the processing is performed with the big astronomical data along with the configuration parameters of each mathematical and processing module, the data mining approach is very useful [16]. Data mining is an analysis step of the "knowledge discovery in databases" (KDD) process [17]. The data mining carries out about the useful information extracting using the intelligent methods from a data set or configuration parameters set to transform it according to the required contracts and protocols and prepare for the further usage in the processing pipeline [18]. In astronomy almost all astronomical images are made by the CCD-camera [19] and can be received for the processing pipeline from the different sources: archives, servers, predefined series of frames, Virtual Observatories [20], clusters, etc. Such big astronomical data is used for the different image processing and machine vision purposes [21], like analyzing, acquiring, pre-processing, processing, and extraction of high-dimensional astronomical information [22] from the pipelines. Such purposes are focused on but not limited to the following tasks: brightness equalization and background alignment [23], object's images detection [24], moving objects detection [25], astrometry of object's image [26], photometry of object's image [27], the estimation of the object's image and motion parameters [28], reference objects cataloging [29], objects recognition [30], matched filtration [31], time series analysis [32], Wavelet coherence analysis [33] and others. In this chapter we presented a description of the different processing pipelines, selected one of the astronomical scientific software based on such processing pipeline, described the implementation of data mining purposes related to the astronomical processing configuration parameters during invocation the image processing pipeline and its implementation in the developed ThresHolds tool. It is specially designed as a part of the CoLiTec project [34] for working with a big amount of astronomical configuration parameters that are used by the different mathematical and processing modules and components. Also, the implementation of the developed Telescope tool, which is designed for mining the big astronomical data from the different archives, parsing the metadata from each astronomical file, and collecting it with the further insertion into the database. Thus, the main aim of this chapter is the development and research of the helpful tool ThresHolds for data mining of astronomical processing configuration parameters. The rest of this chapter is organized as follows.

Section 2 presents details about the processing pipeline of the selected scientific astronomical software called "CoLiTec". In section 3, the authors discuss in detail the development of the ThresHolds tool and its data mining tasks for the processing configuration parameters. Section 4 outlines the real astronomical examples and successful implementation of the mentioned helpful tool. The chapter ends with a conclusion in section 5 and acknowledgements in section 6.

2. Processing pipeline in the CoLiTec software

For our research we have selected astronomical scientific software for detection of the moving objects in a series of CCD-frames called “CoLiTec” (Collection Light Technology), which implements the image processing pipeline. Such software performs almost all astronomical image processing tasks, like filtering [35], brightness equalization [23], background alignment [3], image stacking/segmentation [25], object detection [4], motion detection [5], object astrometry [26], object photometry [27], object’s image and motion parameters estimation [28], machine (computer) vision [23] of the reference objects to be cataloged [31], object recognition [36] time series analysis [32], Wavelet coherence analysis [33], machine learning recognition [30] and others. The modern CoLiTec software was developed using different technologies and approaches for big data processing, data mining, and machine vision. In the astronomy direction, the CoLiTec software is designed to perform the following main stages of machine vision and image processing: pre-processing (astronomical data collection -> worst data rejection -> useful data extraction -> classification -> clustering -> background alignment -> brightness equalization), processing (recognition patterns applying -> machine vision -> object’s image detection -> astrometry -> photometry -> tracks detection), knowledge extraction (astronomical objects to be discovered, tracks parameters for the investigation, light curves of the variable stars). More features of the CoLiTec software are detailed described below [38]:

– processing images with the very wide FOV (<10 degrees^2);

– calibration and cosmetic correction in automated mode using the appropriate calibration master frames and their creation if necessary;

– brightness equalization and background alignment of the images in series using the mathematical inverse median filter;

– rejection of the bad and unclear observations and measurements of the investigated astronomical objects in automated mode;

– fully automatic robust algorithm of the astrometric and photometric reduction of the investigated astronomical objects;

– detection of the faint investigated astronomical moving objects in a series of CCD images with a signal-to-noise ratio (SNR) of more than 2.5 in automated mode;

– detection of the very fast investigated astronomical objects (<40.0 pix./frame) in automated mode;

– detection of the astronomical objects with near-zero apparent motion (from 0.7 pix./frame) in automated mode;

– rejection of the investigated astronomical objects with bad and corrupted measurements in automated mode;

– viewer of the processing results with simple and understandable graphical user interface (GUI) by the LookSky software;

– confirmation of the most interesting astronomical objects at the night of their preliminary discovery;

– multi-threaded processing support;

– multi-cores systems support;

– support of managing the individual threads and processes;

– processing pipeline managed by the On-line Data Analysis System (OLDAS);

– deciding system of the processing results, which allows to adapt the end-user settings and inform the user about the processing results at each stage in the pipeline;

– subject mediator as a major part of the data control in the pipeline during processing.

CoLiTec software realizes the different knowledge discovery in databases and data mining approaches, like pre-processing, clustering, classification, identification, processing, summarization. CoLiTec software is a very complex astronomical system for the big data sets processing, which includes the different features, user-friendly tools for the processing management, results reviewing, integration with online astronomical catalogs and a lot of computational components and modules that are based on the developed mathematical methods [5, 28, 37]. The high level processing pipeline with the developed modules and implemented methods of the CoLiTec software is presented in Fig. 1. Totally, the CoLiTec software consists of more than 30 mathematical and processing modules/components related to the different stages of the image processing that included to the common processing pipeline. Each such module/component has a lot of configuration parameters to be set for the proper image processing and tuning the processing results. All relationships between processing modules and components in the CoLiTec software are based on the predefined contracts as a stable form of the input description for the module processing. Almost all contracts have a fixed structure and description based on the eXtensible Markup Language (XML) [38].

This is a file format with the main purposes to serialize, store, transmit, and reconstruct the different arbitrary data. XML defines a set of rules for encoding documents in a format that is both human-readable and machine-readable.

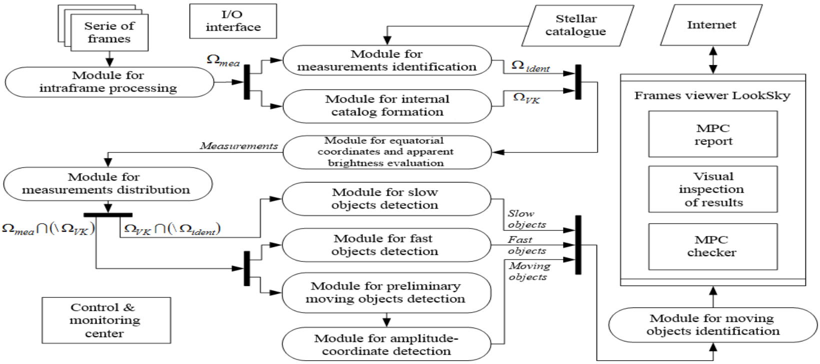


Fig. 1. The high level processing pipeline of the CoLiTec software

The main design goals of XML are to emphasize simplicity, generality, and usability across the processing modules. For the exchanging of different kind of information between two disparate and separate processing modules, they need to agree upon a file format, and the XML standardizes this process. Generally, the XML file is a file with the textual data format with strong Unicode support and a fixed structure. Each XML document has a hierarchical structure and is conceptually interpreted as a tree structure, called the XML tree. Such tree should contain the root element (only one parent element of all other elements), sub elements with attributes and text.

The example of XML file with the astronomical configuration parameters in the CoLiTec software is presented in Fig. 2.


```xml
<CoLiTec>
  <artifactsFilter>
    <IsDetect default="true" type="xs:boolean">true</IsDetect>
    <NegativePixels default="5.0" max="10.0" min="2.0" type="xs:decimal">5.0</NegativePixels>
    <HotPixels default="5.0" max="10.0" min="2.0" type="xs:decimal">5.0</HotPixels>
  </artifactsFilter>
  <astroPhotoMetry>
    <UseOpticalPolynom default="false" type="xs:boolean">false</UseOpticalPolynom>
    <MarksFromFrame default="50" max="100" min="30" type="xs:integer">50</MarksFromFrame>
    <PercentOfMatching default="70" max="80" min="60" type="xs:integer">70</PercentOfMatching>
    <StarsResearch default="0.0" max="1.0" min="0.0" type="xs:decimal">0.0</StarsResearch>
    <LimitOfMagnitude default="25.0" max="25.0" min="15.0" type="xs:decimal">25.0</LimitOfMagnitude>
    <DifferentBetweenAngles default="3" max="10" min="1" type="xs:decimal">3</DifferentBetweenAngles>
    <DistanceTo3DVertex default="0.3" max="0.9" min="0.1" type="xs:decimal">0.3</DistanceTo3DVertex>
    <AstroReductionDirection default="XYtoRADE" list="RADEtoXY,XYtoRADE" type="xs:string">XYtoRADE</AstroReductionDirection>
    <AstrometryReductionDegree default="3" list="1,3,5" type="xs:integer">3</AstrometryReductionDegree>
    <ReductionDegree default="1" list="1,3,5" type="xs:integer">1</ReductionDegree>
    <OpticalPolynomDegree default="5" list="1,3,5" type="xs:integer">5</OpticalPolynomDegree>
    <StarsRejectionThreshold default="0.1" max="10.0" min="0.0" type="xs:decimal">0.1</StarsRejectionThreshold>
    <PairReductionCoefficient default="1.0" max="99.9" min="0.1" type="xs:decimal">1.0</PairReductionCoefficient>
    <FragmentsNumber default="4" max="1000" min="1" type="xs:integer">4</FragmentsNumber>
    <AmplitudeSort default="A2" list="A2,A3" type="xs:string">A2</AmplitudeSort>
    <UseAstrocatalogErrors default="false" type="xs:boolean">false</UseAstrocatalogErrors>
    <DoIC default="false" type="xs:boolean">false</DoIC>
    <DoNotSelectStarsOnFrame default="(0;0-0;0)(0;0-0;0)" type="xs:string">(0;0-0;0)(0;0-0;0)</DoNotSelectStarsOnFrame>
    <CriticalSNR3forMeasurements default="5.0" max="999.9" min="0.1" type="xs:decimal">5.0</CriticalSNR3forMeasurements>
    <MinPixelScale default="1.0" max="180.0" min="1.0" type="xs:decimal">1.0</MinPixelScale>
    <MaxPixelScale default="180.0" max="180.0" min="0.1" type="xs:decimal">180.0</MaxPixelScale>
  </astroPhotoMetry>
  <brightnessFramesAlignment>
    <MedianWindow default="101" max="500" min="50" type="xs:integer">101</MedianWindow>
    <BinningCoefficient default="3" max="16" min="1" type="xs:integer">3</BinningCoefficient>
    <SpecifiedMasterFrames default="false" type="xs:boolean">false</SpecifiedMasterFrames>
    <UseCommonFolder default="false" type="xs:boolean">false</UseCommonFolder>
    <ProcessingRawFolder default="" type="xs:string"></ProcessingRawFolder>
    <UseMasterBias default="false" type="xs:boolean">false</UseMasterBias>
    <ProcessingRawBias default="" type="xs:string"></ProcessingRawBias>
    <UseMasterDark default="false" type="xs:boolean">false</UseMasterDark>
    <ProcessingRawDark default="" type="xs:string"></ProcessingRawDark>
    <UseMasterDarkFlat default="false" type="xs:boolean">false</UseMasterDarkFlat>
    <ProcessingRawDarkFlat default="" type="xs:string"></ProcessingRawDarkFlat>
    <UseMasterFlat default="false" type="xs:boolean">false</UseMasterFlat>
    <ProcessingRawFlat default="" type="xs:string"></ProcessingRawFlat>
    <UseFHF default="true" type="xs:boolean">true</UseFHF>
    <IgnoreTimeFactor default="false" type="xs:boolean">false</IgnoreTimeFactor>
    <ExposureAutoDetermination default="false" type="xs:boolean">false</ExposureAutoDetermination>
  </brightnessFramesAlignment>
```


Fig. 2. The example of XML file with the astronomical configuration parameters in the CoLiTec software

Totally, there are more than 700 astronomical configuration parameters related to the different image processing tasks performed by up to 30 processing modules in the pipeline of CoLiTec software. The structure of such XML configuration file has the restricted rules:

– each element of the XML tree has the required attributes: “default” and “type”;

– “default” and “type” attributes should not be empty;

– “type” attribute should have only fixed values according to the XML Schema Definition (XSD) [39] defined by the World Wide Web Consortium (W3C) [40]: “xs:string”, “xs:boolean”, “xs:integer”, “xs:decimal”;

– each element should contain the text;

– “list” attribute is allowed with coma separator only for element with type “xs:string”;

– “min” and “max” attributes are required only for element with types “xs:integer” and “xs:decimal”;

– “min” and “max” attributes should have the values according to the “type” attribute;

– “node” attribute should contain only the following values: “checkbox”, “textfield”, and “radiobutton”;

– “id”, “label” and “node” attributes are optional, but should not be empty;

– “id” attributes should contain the unique value;

– “label” attribute should contain the text value.

3. Data mining by the ThresHolds tool

Under the research in scope of the CoLiTec project [34] we have developed the ThresHolds tool for the following data mining and processing tasks: mining the astronomical processing configuration parameters; classification, managing and validating of the astronomical processing configuration parameters; visualization for the end user; preparing the batch of required configuration parameters for the appropriate stage in processing pipeline of the big astronomical data from the different storages and archives.

The ThresHolds tool realized the different data mining tasks, like receiving, storing, selecting, preprocessing, transforming, useful data extraction, classification, and knowledge discovery in databases (KDD) [17]. The following stack of different technologies were used for the ThresHolds tool development: Java programming language [41]; JavaFX technology for graphical user interface (GUI); Maven for compilation and building; Java Development Kit (JDK) [42]; Open JDK; XML [38]. As a JDK the developers selected the Open JDK as an open-source and free Java platform including the Java machine, which is supported by the Java community instead of the Oracle JDK. The Open JDK plays the role as a platform for developing the client applications for desktop, laptop, and tablet PCs and is a cross-platform, so it supports the different operational systems (OSs), like Windows, Linux, MacOS. The ThresHolds tool has a GUI for visualization of the processing configuration parameters for the end user and the additional useful features described below. The ThresHolds tool is designed for the full integration with astronomical image processing pipeline of the CoLiTec software and performs the following data mining tasks:

– recurrency searching for the configuration XML files;

– visualization by the dynamically creating the GUI according to the data in the configuration XML file;

– managing the astronomical configuration parameters by the end user using the GUI (see some examples of different sections in Fig. 3 - 6);

– mining the astronomical processing configuration parameters from the modified XML file;

– classifying the astronomical configuration parameters according to the different image processing tasks and stages in the processing pipeline;

– validating the astronomical configuration parameters according to the restricted rules;

– dynamically generating the batch of required parameters for the appropriate image processing tasks and stages in the processing pipeline.

The task panel (see Fig. 7) in the ThresHolds tool allows selecting appropriate processing task. According to the selected processing task, the user can perform accurate adjustment of all parameters that are available in the tree menu.

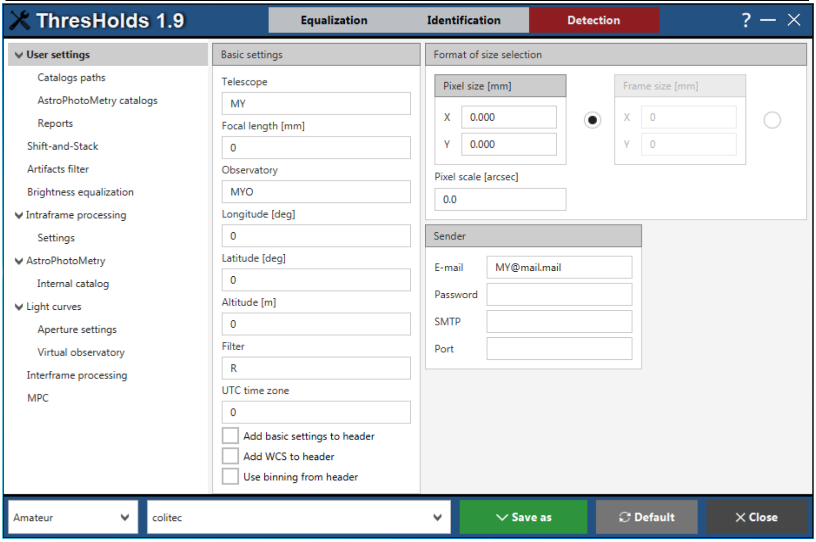


Fig. 3. The "Basic settings" section of the ThresHolds tool

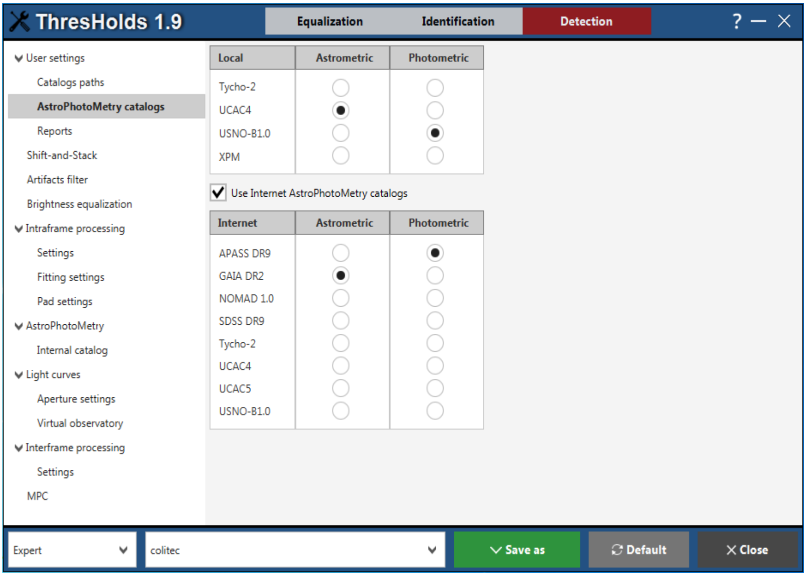


Fig. 4. The "AstroPhotoMetry catalogs" section of the ThresHolds tool

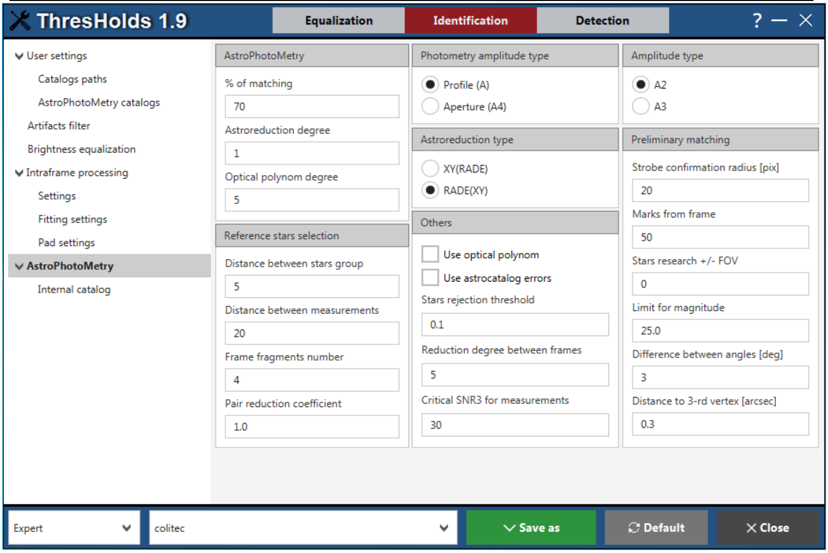


Fig. 5. The "AstroPhotoMetry" section of the ThresHolds tool

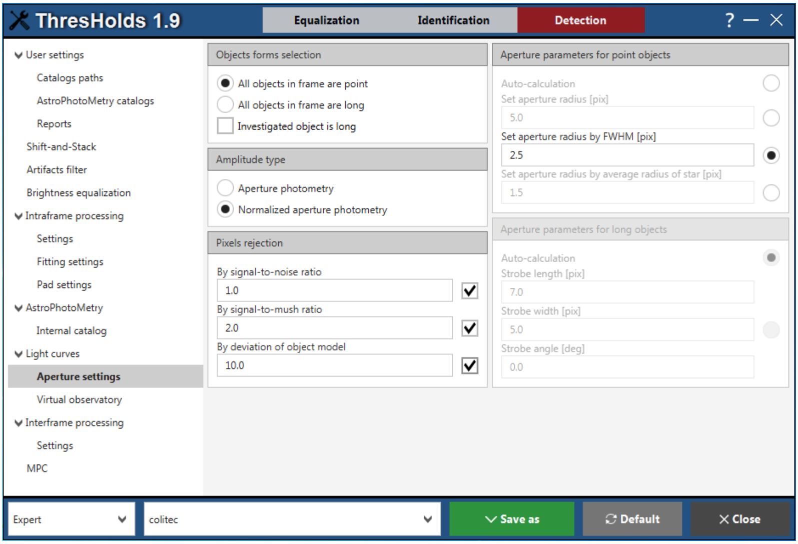


Fig. 6. The "Aperture settings" section of the ThresHolds tool

The following processing tasks are available for the user: equalization, identification, and detection. For the "*Equalization*" processing task the following sections will be available: User settings, Artifacts filter and Brightness equalization. For the "*Identification*" processing task the following sections will be available: User settings, Artifacts filter, Brightness equalization, Intraframe processing and AstroPhotoMetry.

For the "*Detection*" processing task all sections will be available.

Fig. 7. *The task panel for selecting appropriate processing task in the ThresHolds tool*

Also, the following expert levels are available for the user: amateur, expert. According to the selected level some additional sections / fields will be available for the more accurate and professional adjustment of parameters for processing. For example, see the difference between "*Interframe processing*" section in the ThresHolds tool with "*Amateur*" expert level (Fig. 8) and "*Expert*" level (Fig. 9).

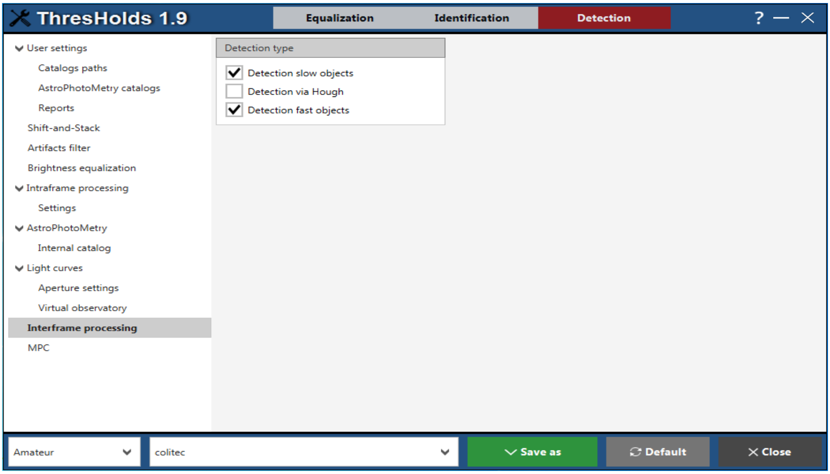


Fig. 8. The "Interframe processing" section of the ThresHolds tool with "Amateur" expert level

Such dynamic changing of the user interface (UI) visualization is also related to the on-fly data mining of the configuration parameters and preparing a batch of the UI controls.

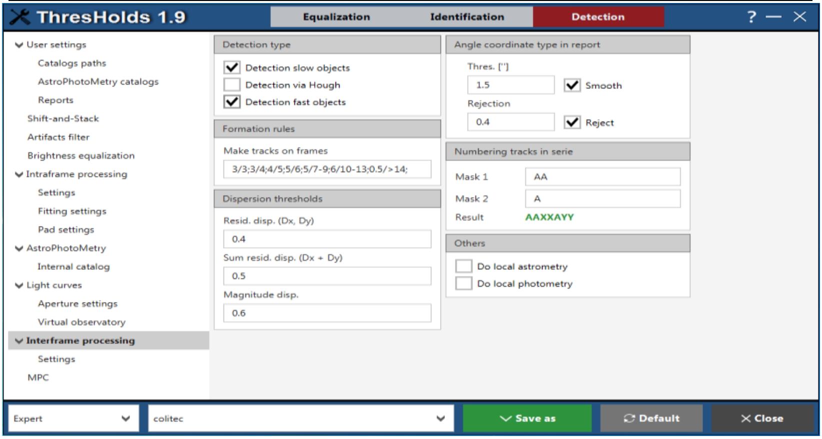


Fig. 9. The "Interframe processing" section of the ThresHolds tool with "Expert" level

For example, to perform the brightness equalization (Fig. 10) by the inverse median filter in the processing pipeline, the ThresHolds tool prepares the batch of the following required parameters for processing:

– "*Binning coefficient*" to perform the median filtering with binning image. Such result is de-binning and subtracted from the original image.

– "*Size of filter mask*" as a width and height of the square median filter mask m. The recommended value of it is determined by the following equation:

$$m \geq \sqrt{3\pi}R, \qquad (1)$$

where R is an image radius of the brightest object by $2\hat{\sigma}_{noise}$ level of background;

$\hat{\sigma}_{noise}$ is a root-mean-square (RMS) evaluation of the background brightness.

– "*Do image inversion*" – activates the image inversion function during transformation.

– "*Make crop*" – activates the crop creation function. It defines by the coordinates of the upper left vertex, the width and height.

– Crop parameters can be set in "X", "Y", "Width" and "Height" fields.

– "*Output mask*" – mask for file names of the processed images.

– “*Process frames by RGB channels*” allows processing a color image by channels. It does not affect on filtering of gray images and FITS files [43];

– “*Path to input file(s)*” – path to the raw files.

– “*Path to output folder*”– path to the output folder of processed files.

– Option for creating and using of the master-frames.

– Paths of the created master-frames will be inserted to the appropriated fields: a. For Bias – “*Path to Master-Bias*”; b.For Dark – “*Path to Master-Dark*”; c. For DarkFlat – “*Path to Master-DarkFlat*”; d.For Flat – “*Path to Master-Flat*”:

– “*Mask after subtraction*” – a prefix to the name of the target (Light) frames after subtraction operation;

– “*Mask after division*” – a prefix to the name of the target (Light) frames after division operation;

– “*Pixels rejection RMS*” – coefficient of the pixel’s rejection in an operation of master-frames creation.

An example of the XML configuration file for the mathematical module of the inverse median filtering formed but the ThresHolds tool as a batch of the required parameters for processing is presented in Fig. 11.

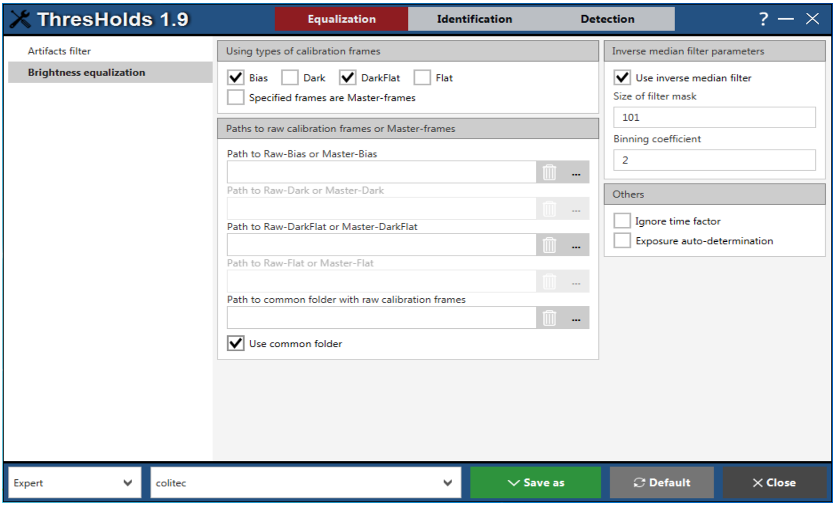


Fig. 10. The “Brightness equalization” section of the ThresHolds tool

```xml
<FhfSettings>
  <highpassFilter>
    <IsHighpassFilter>0</IsHighpassFilter>
    <FullNameFitInFilter>E:\Ubuntu15\refhflinux\flat-002R.fit</FullNameFitInFilter>
    <FullNameFitOutFilter>E:\Ubuntu15\refhflinux\_fhf99_flat-002R.fit</FullNameFitOutFilter>
    <DayFit>0</DayFit>
    <RadiusFilter>100</RadiusFilter>
    <NumberIter>12</NumberIter>
    <KSkoFilter>0</KSkoFilter>
  </highpassFilter>
  <crop>
    <IsCrop>0</IsCrop>
    <FullNameFitInCrop>E:\Ubuntu15\refhflinux\flat-002R.fit</FullNameFitInCrop>
    <PathFitOutCrop>E:\Ubuntu15\refhflinux\flat-002R_crop\</PathFitOutCrop>
    <SizeUp>200</SizeUp>
    <SizeDown>200</SizeDown>
    <SizeLeft>200</SizeLeft>
    <SizeRight>200</SizeRight>
    <Invert>0</Invert>
    <Crop>3</Crop>
    <CoordinateX>1000</CoordinateX>
    <CoordinateY>2000</CoordinateY>
    <WidthCrop>1000</WidthCrop>
    <HeightCrop>2000</HeightCrop>
  </crop>
  <inversMedianFilter>
    <IsInversMedianFilter>1</IsInversMedianFilter>
    <FullNameFitInInversMedianFilter>e:\Ubuntu15\refhflinux\RXJ1803 -114R.fit</FullNameFitInInversMedianFilter>
    <FullNameFitOutInversMedianFilter>e:\Ubuntu15\refhflinux\RXJ1803 -114R_w51_b2.fit</FullNameFitOutInversMedianFilter>
    <InversMedianWindow>51</InversMedianWindow>
    <ColorOrBw>0</ColorOrBw>
    <SmallOrBig>0</SmallOrBig>
    <Binning>2</Binning>
  </inversMedianFilter>
</FhfSettings>
```


Fig. 11. The formed XML configuration file for the mathematical module of the inverse median filtering

As mentioned above in Fig. 1, the ThresHolds tool is included into the high level processing pipeline of the CoLiTec software in automated and autonomous mode. Each such mentioned module and submodule has its own set of configuration parameters, and the main data mining task of the ThresHolds tool is to collect appropriate data and send it to the processing pipeline. In this case, it is a realization of the subject mediator approach.

The ThresHolds tool operates with astronomical configuration parameters, which are equal to the scientific constants or variables of the different mathematical methods and algorithms for the astronomical image processing.

The developed algorithm for data mining tasks realized in the ThresHolds tool contains the following main steps.

1. Processing pipeline selects the next module, which is predefined in a sequence.

2. Such processing module has the appropriate template or skeleton for the configuration file with processing parameters required for it.

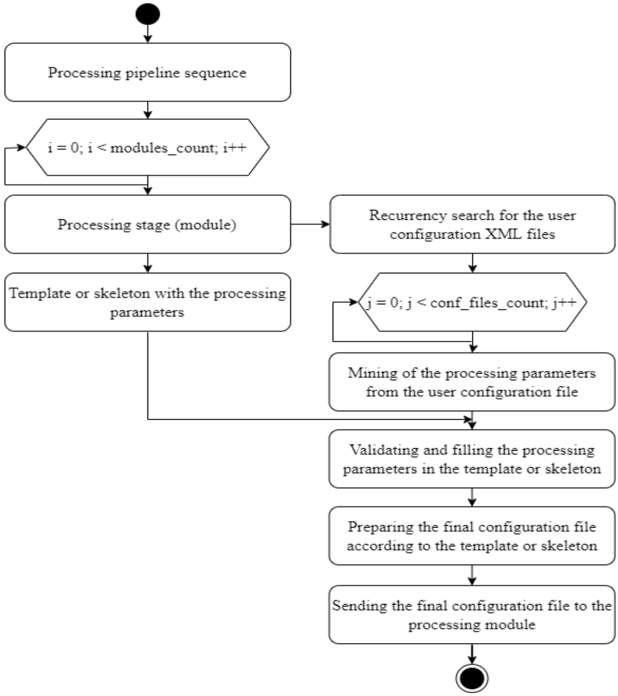


Fig. 12. The high-level diagram of the developed algorithm for data mining realized in the ThresHolds tool

3. Classification of the astronomical configuration parameters according to the different image processing tasks and stages in the processing pipeline.

4. Recurrency search for the different user configuration XML files prepared by user before processing starts.

5. Processing parameters mining from the selected user configuration XML file according to the template or skeleton.

6. Validating of the astronomical configuration parameters according to the restricted rules;

7. Dynamically generating the batch of required parameters according to the template or skeleton for the appropriate image processing tasks and stages in the processing pipeline.

8. Processing pipeline selects the next processing module, which is predefined in a sequence.

4. Practical implementation with the real astronomical examples

The ThresHolds tool in scope of the CoLiTec software were installed in the different observatories (Mayaki Astronomical Observatory [44, 45], ISON-NM and ISON-Kislovodsk observatories [46], Vihorlat Observatory [2, 23]), astronomical archives [20], and Ukrainian Virtual Observatory (UkrVO) [47]. An example of data mining of the astronomical processing configuration parameters by the ThresHolds tool integrated into the processing pipeline is presented in Fig. 13.

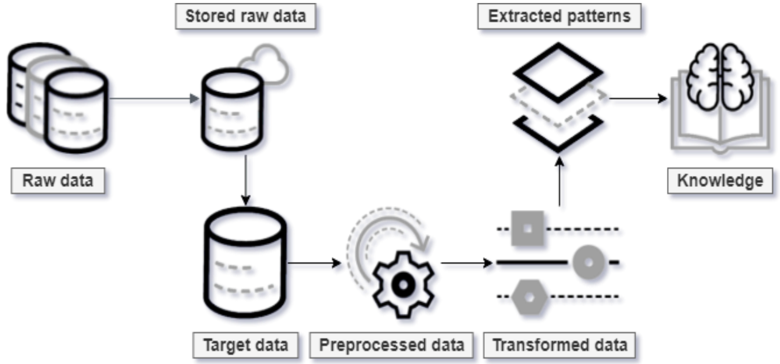


Fig. 13. Example of data mining of the astronomical processing configuration parameters by the ThresHolds tool integrated into the processing pipeline

More detailed information about the observatories, telescopes equipped by the different CCD-cameras as well as their parameters are provided below. The Mayaki observing station of "Astronomical Observatory" Research Institute of I. I. Mechnikov Odessa National University has the 0.48 m AZT-3 telescope – reflector with focal length 2025 mm and CCD-camera Sony ICX429ALL (resolution 795×596). The observatory "ISON-NM observatory" has the 0.4 m SANTEL-400AN telescope with CCD-camera FLI ML09000-65 (3056×3056 pixels, 12 microns). The observatory "ISON-Kislovodsk" has the 19.2 cm wide-field GENON (VT-78) telescope with CCD-camera FLI ML09000-65 (4008×2672 pixels, 9 microns).

The observatory "Vihorlat Observatory in Humenne" has the Vihorlat National Telescope (VNT) – Kassegren telescope with 1 m main mirror with focal length 8925 mm and CCD-camera FLI PL1001E (512×512 pixels).

The Vihorlat Observatory also has the Celestron C11 telescope – Schmidt-Cassegrain telescope with 28 cm main mirror with focal length 3060 mm and CCD-camera G2-1600 (resolution 768×512 pixels).

The data mining of astronomical processing configuration parameters by the ThresHolds tool was performed during the processing of up to one million astronomical configuration files both archived and original formed from the different telescopes of the various observatories.

5. Conclusion

The ThresHolds tool with the realization of the developed algorithm for data mining [48] purposes related to the astronomical processing configuration parameters during invocation the image processing pipeline was developed. The tool is implemented using the Java programming language and JavaFX technology. The advantages of such selected programming language is an open-source libraries and cross platform approach.

The ThresHolds tool was developed for the full integration with the astronomical image processing pipelines with a GUI for visualization of more than 700 processing configuration parameters.

The main goals of the ThresHolds tool are the mining of astronomical processing configuration parameters, their classification, managing and validation. Also, visualization for end user is available as well as preparation the batch of required parameters for the appropriate stage in processing pipeline of the big astronomical data from the different storages and archives.

Research showed that using the developed algorithm for data mining purposes the ThresHolds tool execution time speeds up in comparison to the common algorithms for the data processing without optimization and mediator subject approach.

In total, based on the statistical [49] experiments result is more than 30% of speeding up of the processing time of the whole pipeline and the whole sequence of the astronomical scientific data processing. The ThresHolds tool was developed as a part of the CoLiTec software. It was tested for several years after successful installation in scope of the astronomical image processing pipelines on the different observatories.

6. Acknowledgements

The authors thank all observatories for their collaboration, providing an access to their archives with astronomical data and possibility to install the ThresHolds tool in scope of the CoLiTec software as a main astronomical processing pipeline to conduct the current research and test the developed software.

7. References

[1] C. A. Politsch, et al., Trend filtering–I. A modern statistical tool for time-domain astronomy and astronomical spectroscopy, MNRAS, vol. 492, issue 3, 2020, pp. 4005-4018. DOI: 10.1093/mnras/staa106.

[2] Š. Parimucha, et al., CoLiTecVS – A new tool for an automated reduction of photometric observations, Contributions of the Astronomical Observatory Skalnate Pleso, vol. 49, issue 2, 2019, pp. 151-153.

[3] E. Corsaro, et al., Spin alignment of stars in old open clusters, Nature Astronomy 1 4 (2017) 1-6, DOI: 10.1038/s41550-017-0064.

[4] L. Mykhailova, et al., Method of maximum likelihood estimation of compact group objects location on CCD-frame, Eastern-European Journal of Enterprise Technologies 5 4 (2014) 16-22. DOI:10.15587/1729-4061.2014.28028.

[5] S. Khlamov, et al., Development of computational method for detection of the object's near-zero apparent motion on the series of CCD–frames, Eastern-European Journal of Enterprise Technologies 2 80 (2016) 41-48. DOI: 10.15587/1729-4061.2016.65999.

[6] R. Klette, Concise computer vision, Springer, London, 2014.

[7] M. Dadkhah, et al., Methodology of wavelet analysis in research of dynamics of phishing attacks, International Journal of Advanced Intelligence Paradigms 12 4 (2019) 220-238. DOI: 10.1504/IJAIP.2019.098561.

[8] G. K. Adam, et al., Embedded microcontroller with a CCD camera as a digital lighting control system, Electronics 8 1 (2019). DOI: 10.3390/electronics8010033.

[9] A. Tantsiura, et al., Evaluation of the potential accuracy of correlation extreme navigation systems of low-altitude mobile robots, International Journal of Advanced Trends in Computer Science and Engineering 8 5 (2019) 2161–2166. DOI: 10.30534/ijatcse/2019/47852019.

[10] N. Yeromina, V. Tarshyn, S. Petrov, et al., Method of reference image selection to provide high-speed aircraft navigation under conditions of rapid change of flight trajectory, International Journal of Advanced Technology and Engineering Exploration 8 85 (2021) 1621–1638. DOI: 10.19101/IJATEE.2021.874814.

[11] M. Ivanov, et al., Effective informational entropy reduction in multi-robot systems based on real-time TVS, in: IEEE International Symposium on Industrial Electronics, 2019, pp. 1162–1167. DOI: 10.1109/ISIE.2019.8781209.

[12] V. Akhmetov, et al., Cloud computing analysis of Indian ASAT test on March 27, 2019, in: Proceedings of the 2019 IEEE International Scientific-Practical Conference: Problems of Infocommunications Science and Technology, PIC S and T 2019, 2019, pp. 315–318. DOI: 10.1109/PICST47496.2019.9061243.

[13] D. Oszkiewicz, V. Troianskyi, et al., Spins and shapes of basaltic asteroids and the missing mantle problem. Icarus, 2023. DOI: 10.1016/j.icarus.2023.115520.
[14] A. Jangda, U. Bondhugula, An effective fusion and tile size model for optimizing image processing pipelines, ACM SIGPLAN Notices, 53 1 (2018) 261-275. DOI: 10.1145/3200691.3178507.
[15] J. K. Li, K.-L. Ma, P5: Portable progressive parallel processing pipelines for interactive data analysis and visualization, in: IEEE transactions on visualization and computer graphics, 26 1 (2019) 1151-1160. DOI: 10.1109/TVCG.2019.2934537.
[16] S. Khlamov, et al., Data Mining of the Astronomical Images by the CoLiTec Software, in: CEUR Workshop Proceedings, vol. 3171, 2022, pp. 1043–1055.
[17] Q. Luo, Advancing knowledge discovery and data mining, in: First International Workshop on Knowledge Discovery and Data Mining (WKDD 2008), IEEE, 2008, pp. 3-5. DOI: 10.1109/WKDD.2008.153.
[18] Y. Zhang, Y. Zhao, C. Cui, Data mining and knowledge discovery in database of astronomy, Progress in Astronomy 20 4 (2002) 312-323.
[19] H. F. Zhang, C. Chen, J. Wang, et al., Scientific CCD camera for the CSTAR2 telescope in Antarctica, Journal of Astronomical Telescopes, Instruments, and Systems 5 3 (2019). DOI: 10.1117/1.JATIS.5.3.036002.
[20] I. B. Vavilova, et al., Surveys, catalogues, databases, and archives of astronomical data, Knowledge Discovery in Big Data from Astronomy and Earth Observation, Astrogeoinformatics (2020) 57-102. DOI: 10.1016/B978-0-12-819154-5.00015-1.
[21] C. Steger, M. Ulrich, C. Wiedemann, Machine vision algorithms and applications, John Wiley & Sons, 2018.
[22] V. Akhmetov, et al., New approach for pixelization of big astronomical data for machine vision purpose, in: IEEE International Symposium on Industrial Electronics, 2019, pp. 1706–1710. DOI: 10.1109/ISIE.2019.8781270.
[23] I. Kudzej, et al., CoLiTecVS – A new tool for the automated reduction of photometric observations, Astronomische Nachrichten 340 3 (2019) 68-70. DOI: 10.1002/asna.201913562.
[24] V. Savanevych, et al., Formation of a typical form of an object image in a series of digital frames, Eastern-European Journal of Enterprise Technologies 6 120 (2022) 51–59. DOI: 10.15587/1729-4061.2022.266988.
[25] R. Gonzalez, R. Woods, Digital image processing, 4th edition, NY: Pearson, 2018.
[26] V. Akhmetov, et al., Astrometric reduction of the wide-field images, Advances in Intelligent Systems and Computing, 1080 (2020) 896–909. DOI: 10.1007/978-3-030-33695-0_58.
[27] V. Savanevych, et al., CoLiTecVS software for the automated reduction of photometric observations in CCD-frames, Astronomy and Computing 40 (2022) 15. DOI: 10.1016/j.ascom.2022.100605.
[28] V. Savanevych, et al., A method of immediate detection of objects with a near-zero apparent motion in series of CCD-frames. Astronomy & Astrophysics 609 A54 (2018). DOI: 10.1051/0004-6361/201630323.

[29]V. Savanevych, et al., Selection of the reference stars for astrometric reduction of CCD-frames, Advances in Intelligent Systems and Computing, 1080 (2020) 881–895. DOI: 10.1007/978-3-030-33695-0_57.

[30]L. Kirichenko, P. Zinchenko, T. Radivilova, Classification of time realizations using machine learning recognition of recurrence plots, Advances in Intelligent Systems and Computing 1246 (2021) 687–696. DOI: 10.1007/978-3-030-54215-3_44.

[31]S. Khlamov, et al., Development of computational method for matched filtration with analytic profile of the blurred digital image, Eastern-European Journal of Enterprise Technologies 5 119 (2022) 24–32. DOI: 10.15587/1729-4061.2022.265309.

[32]L. Kirichenko, A.S.A. Alghawli, T. Radivilova, Generalized approach to analysis of multifractal properties from short time series, International Journal of Advanced Computer Science and Applications 11 5 (2020) 183–198. DOI: 10.14569/IJACSA.2020.0110527.

[33]V. Baranova, et al., Wavelet Coherence as a Tool for Studying of Economic Dynamics in Infocommunication Systems, in: IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T), 2019, pp. 336-340. DOI: 10.1109/PICST47496.2019.9061301.

[34]S. Khlamov, V. Savanevych, Big astronomical datasets and discovery of new celestial bodies in the Solar System in automated mode by the CoLiTec software, Knowledge Discovery in Big Data from Astronomy and Earth Observation, Astrogeoinformatics: Elsevier, 2020, pp. 331–345. DOI: 10.1016/B978-0-12-819154-5.00030-8.

[35]S. Khlamov, V. Savanevych, V. Vlasenko, O. Briukhovetskyi, T. Trunova, I. Levykin, V. Shvedun, I. Tabakova, Development of the matched filtration of a blurred digital image using its typical form. Eastern-European Journal of Enterprise Technologies 9 121 (2023) 62-71. DOI: 10.15587/1729-4061.2023.273674.

[36]S. Khlamov, I. Tabakova, T. Trunova, Recognition of the astronomical images using the Sobel filter, in: Proceedings of the International Conference on Systems, Signals, and Image Processing, 2022. DOI: 10.1109/IWSSIP55020.2022.9854425.

[37]V. Savanevych, et al., Mathematical methods for an accurate navigation of the robotic telescopes, Mathematics 11 10 (2023). DOI: 10.3390/math11102246.

[38]M. Wang, J. Wang, K. Guo, Extensible markup language keywords search based on security access control, International Journal of Grid and Utility Computing 9 1 (2018) 43-50. DOI: 10.1504/IJGUC.2018.090227.

[39]S. Guo, H. Xia, G. Xiang, Research on the translation from XSD to JSON schema, in: IEEE 9th International Conference on Communication Software and Networks (ICCSN), 2017, pp. 1393-1396. DOI: 10.1109/ICCSN.2017.8230338.

[40]T. Brooks, World Wide Web Consortium (W3C), Encyclopedia of Library and Information Sciences, CRC Press, 2017.

[41]G. O’Regan, Java programming language, The Innovation in Computing Companion. Springer, Cham, 2018. DOI: 10.1007/978-3-030-02619-6_35.

[42]D. Winnie, Setting Up the Java JDK and IntelliJ, Essential Java for AP CompSci. Apress, Berkeley, CA, 2021. DOI: 10.1007/978-1-4842-6183-5_2.

[43] W. D. Pence, L. Chiappetti, C. G. Page, R. A. Shaw, and E. Stobie, Definition of the flexible image transport system (fits), version 3.0, Astronomy & Astrophysics 524 A42 (2010). DOI: 10.1051/0004-6361/201015362.

[44] V. Troianskyi, P. Kankiewicz, D. Oszkiewicz, Dynamical evolution of basaltic asteroids outside the Vesta family in the inner main belt, Astronomy and Astrophysics 672 A97 (2023). DOI: 10.1051/0004-6361/202245678.

[45] S. Kashuba, et al., The Simeiz plate collection of the ODESSA astronomical observatory, in: 11th Bulgarian-Serbian Astronomical Conference, 2018, pp. 207-216.

[46] I. Molotov, et al., ISON worldwide scientific optical network, in: Proceedings of the Fifth European Conference on Space Debris, ESA SP-672, 7, 2009.

[47] I. B. Vavilova, et al., Astrometry and photometry of asteroids from the UkrVO database of astroplates, in: proceedings of the IAU, vol. 15, issue S364, 2021, pp. 239-245. DOI: 10.1017/S1743921322000047.

[48] M. Brescia, G. Longo, Astroinformatics, data mining and the future of astronomical research, in: Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, vol. 720, 2012, pp. 92-94. DOI: 10.1016/j.nima.2012.12.027.

[49] V. Shvedun, S. Khlamov, Statistical modelling for determination of perspective number of advertising legislation violations, Actual Problems of Economics, 184 10 (2016) 389-396.

УДК 004.853

КОРИСНИЙ ПОРОГОВИЙ ІНСТРУМЕНТ ДЛЯ МІНІНГУ ПАРАМЕТРІВ КОНФІГУРАЦІЇ АСТРОНОМІЧНОЇ ОБРОБКИ

Ph.D. Хламов С.

Харківський національний університет радіоелектроніки, Україна
E-mail: sergii.khlamov@gmail.com

Анотація. *Глава присвячена реалізації цілей інтелектуального аналізу даних, пов'язаних з параметрами конфігурації астрономічної обробкипід час виклику конвеєра обробки зображень. Кожен модуль обробкиконвеєрі має безліч параметрів конфігурації, які дозволяють налаштувати процес обробки, а також підвищити точність і швидкість обчислень під час обробки. Через велику кількість таких параметрів конфігурації в складному конвеєрі наукового програмного забезпечення підхід інтелектуального аналізу даних є дуже корисним і продуктивним. Для нашого дослідження ми вибрали наукове програмне забезпечення для виявлення рухомих об'єктів у серії ПЗЗ-кадрів під назвою "CoLiTec". Таке програмне забезпечення виконує багато різних завдань обробки зображень, таких як фільтрація, вирівнювання фону, виявлення об'єктів, астрометрія, фотометрія, виявлення руху тощо. Програмне забезпечення CoLiTec складається з понад 30 математичних модулів і модулів обробки, пов'язаних з різнимиетапами обробки зображень, де кожен із них має багато параметрів конфігурації, які потрібно встановити. Отже, щоб вирішити проблеми з керуванням параметрами конфігурації астрономічної обробки, розробники вирішили створити новий інструмент для інтелектуального аналізу таких параметрів під назвою «ThresHolds». Він реалізований у вигляді програмного забезпечення з графічним інтерфейсом користувача (GUI) з використанням мови програмування Java, технології JavaFX. Основна мета інструменту ThresHolds - класифікувати параметри конфігурації, керувати ними та перевіряти їх, візуалізувати для кінцевого користувача та підготувати пакет необхідних параметрів для відповідного етапу конвеєра обробки. Інструмент ThresHolds у рамках програмного забезпечення CoLiTec було успішно встановлено як основний конвеєр обробки астрономічних зображень у різних обсерваторіях.*

Ключові слова: *інтелектуальний аналіз даних, великі дані, конвеєр обробки, потік даних, параметри конфігурації, обробка зображень, машинне бачення, CoLiTec*

Section 5. Intellectual models and knowledge engineering technologies

UDC 004:37:001:62

INTELLIGENT MONITORING OF THE TECHNICAL CONDITION OF COMPLEX SYSTEMS

Vychuzhanin A.

ORCID: 0000-0001-8779-2503
National University "Odessa Polytechnic"
e-mail: v.v.vychuzhanin@op.edu.ua

Abstract. *The relevance of the topic is due to the need to make decisions to ensure the reliability of elements and assemblies of complex technical systems with insufficient information about their technical condition. The effectiveness of solving the problem lies in the use of information technologies and artificial intelligence methods, in particular expert systems. The article presents an intelligent system that operates using a developed model for assessing and predicting the risk of failure of components of a complex technical system using the example of a ship's power plant. The construction of a model taking into account the hierarchical levels of subsystems (components), intersystem (interelement) connections of an intelligent system is based on the use of a priori information about failures of components of complex technical systems. The model connects the types of technical condition of components and diagnostic features of systems in the form of the risk of their failures. The use of posterior inference in Bayesian belief networks allows us to determine the risk of failure of system components, taking into account incoming diagnostic information and information about component failures. An intelligent decision support system has been developed that allows assessing the risk of failure of elements and components of complex technical systems using elements of artificial intelligence. The proposed decision support system contains: a database; a knowledge base with methods for calculating reliability indicators (probabilities and risks of failures) and a set of decision rules for selecting appropriate decision-making methods; results of determining the probabilities and risks of failure of elements and assemblies of complex technical systems with their ranking; intellectualization model for assessing the technical condition of elements and assemblies. The proposed algorithm for the functioning of a decision support system implements the task of automating the process of assessing the technical condition of complex systems. The results of studies of a model for assessing and predicting the risk of failure of components of a complex technical system confirmed the possibility of predicting the risk of failure of components and the system as a whole. The use of the proposed decision support system for assessing the technical condition of complex systems will improve the reliability of technical systems with insufficient information about their technical condition.*

Keywords: *complex technical systems, reliability, risk of failure, diagnostics, decision support, prediction, intelligent system, Bayesian belief network*

1. Introduction

The complexity of the composition and the increase in the number of technical systems installed at various facilities lead to an increase in the intensity of their failures. As a result, there is a need to repair the equipment of the systems, which leads to its downtime.

When designing, manufacturing and operating complex technical systems (CTS), reliability is ensured by methods and means specific to each stage of the "life cycle" of systems.

The operational reliability of the restored CTS and their components is effectively achieved by the strategy of operating systems with technical condition monitoring based on technical diagnostic systems [1, 2, 3, 4, 5]. The reduction of failures and man-made risks during the operation of CTS is facilitated by the prediction of their technical condition based on diagnostics.

Currently, the volume of implementation of automation, digitalization and artificial intelligence technologies in various industries continues to grow. For example, in accordance with the requirements of the Register of Maritime Navigation, all modern ships must be equipped with automation systems for technical means using digital technologies, as well as artificial intelligence technologies [2, 6, 7, 8, 9, 10].

Such systems should constantly monitor the components of the ship's CTS, analyze trends in changing the operating modes of the equipment of the systems, perform emergency transfers and provide decision support. To implement such a technology, appropriate algorithmic and software tools are needed to provide diagnostics, forecasting the technical state of systems, and support for decision-making that is adequate to the goal.

The diagnostic algorithms used, as a rule, are based on the tolerance control of individual diagnostic parameters. At the same time, the volume of measuring and diagnostic information, the number of connections, dependencies of diagnostic features and types of technical states of systems can be significant. In theory, engineering practice, various methods are used to assess the risk of failure of CTS components.

An example of the application of risk theory is the logical development of a probabilistic approach for assessing the risk of failures [11, 12]. With a probabilistic approach, the level of reliability is selected depending on the possible consequences of damage (failure) of system components. In this regard, the assessment of the risk of CTS failures lies in the unacceptable probability of their damage.

However, the negative consequences of a failure in systems are often taken into account intuitively, implicitly, by taking certain values of the probability of failure-free operation or the safety factor of system components.

In artificial intelligence, various models of knowledge representation are actively developing. Bayesian belief networks (BBN) are a promising mathematical tool that can be used, in relation to diagnostic tasks, to take into account both the causal relationship between the types of CTS technical condition and diagnostic features, and the arrival of new information in the form of statistical data or predictive estimates.

Bayesian networks allow combining a priori (initial) knowledge about an object with experimental data to obtain an a posteriori estimate [13, 14].

Forecasting the state of CTS plays an important role in planning their operation. It is assumed that the actual technical condition of an object can be assessed by the results of monitoring its parameters, and predicting their changes allows the object to be operated until signs of a dangerous decrease in reliability appear.

There are efficient algorithms and forecasting methods. Artificial intelligence models, in particular, neural networks, are being actively developed to solve forecasting problems [15, 16]. However, the main problem for the productive operation of a neural network is the need for a significant amount of statistical data, which is difficult to obtain in real conditions due to a number of reasons (high cost of the systems under study, high costs for testing, limited time, etc.). The lack of a clear understanding in the choice of neural network architecture for solving various types of problems (pattern recognition, approximation, prediction, etc.) and areas of application also complicates their application.

The conceptual basis for the intellectualization of the solution of interrelated problems of diagnostics, forecasting and decision support is traditional for the class of unstructured and poorly formalized tasks: the impossibility of obtaining complete and objective information for making adequate decisions and the resulting need to involve informal (subjective, heuristic) information; the presence of uncertainty in the initial data, as well as the presence of ambiguity (multiple options) in the process of finding a solution; the need to develop and justify the desired solutions to the problem inconditions of strict time constraints, which are determined by the course of controlled processes; the need to correct and introduce additional information into the process of finding solutions, the interactive (dialogue, human-machine) nature of the logical inference of solutions.

Taking these factors into account forces us to abandon traditional algorithmic methods and models of decision-making and management and move on to intelligent technologies. Combined with the tasks of diagnosing and predicting, the task of modeling the behavior of the CTS acts as a source of data on the state of the object at the stages of system testing.

Thus, during the operation of CTS, an urgent task remains the improvement of methods and models aimed at accurate and prompt assessments, management of the risk of failures of CTS components.

2. Objective and objectives of the study

To assess the reliability of CTS, various methods have been developed and used, often based on the methods of probability theory and mathematical statistics, which makes it possible to automate the process of assessing the reliability of elements and components of complex systems.

However, the stages associated with supporting decisions made to ensure reliability based on the results of its assessment for CTS, in particular ship systems, are often not automated.

As a result, the quality of decisions made to ensure the reliability indicators of such CTS significantly depends on the qualifications of the personnel operating the system [8, 12, 16].

Evolution in information processing leads to the actualization of the task of not only automating the process of assessing the reliability of elements and components of complex systems, but also to the transfer of part of the intellectual sphere of human activity to the sphere of automation of making and supporting management decisions in the field of ensuring the reliability of CTS.

The creation of intelligent decision support systems (IDSS), in the context of progress in the field of information systems and technologies, find significant application in solving complex, difficult to formalize problems, in particular, diagnosing the reliability of CTS. Distinctive features of problems that are difficult to formalize are the incomplete amount of initial data of the problem being solved, inaccuracy, heterogeneity, and significant computational complexity [17, 18].

The purpose of the study is to ensure the reliability of CTS elements and components during operation based on the use of an intelligent decision support system for assessing their technical condition.

The objectives of the study are to develop a IDSS with insufficient information for assessing the technical characteristics of complex systems.

3. Analysis of the operating principle of the IDSS

Intelligent assessment of TC is a process that includes monitoring, diagnostics and, as a result, evaluation of the vehicle while simultaneously working with knowledge and large amounts of information.

This problem can be solved by using an expert decision support system. Decision support system is a computer system that allows the user to solve professional problems based on the use of databases, knowledge and models, by providing conclusions, recommendations, and assessments of possible alternative solutions to the problem. That is, IDSS helps the user solve a complex problem automatically [19].

In general, IDSS are information expert systems. Expert systems used to assess the reliability of CTS elements and assemblies are recommended to be built on the basis of artificial intelligence. This will make it possible to make management decisions in an automated mode, taking into account the specific tasks of monitoring and diagnosing the CTS.The implementation of the IDSS should be based on the use of research results on the model of a specific operating CTS.

They make it possible to determine the probabilities and risks of failure of CTS elements and assemblies. Similar models can be used in the development of IDSS to assess the technical condition of complex systems.

Such systems solve problems: choosing the best solution from many possible ones - optimization; ordering possible solutions according to preferences - ranking. In both problems, the first and most fundamental point is the selection of a set of criteria on the basis of which alternative solutions are evaluated and compared.

4. Main part

A IDSS is proposed to evaluate the CTS TC. In such a system, in contrast to classical artificial intelligence systems, the theory of decision making is applied instead of attempts to “take into account uncertainty” using production rules of the form “IF.

For the practical implementation and operation of IDSS, it is necessary to link the developed models to an expert system containing calculated, experimental, and also data acquired by experts during the operation of the CTS.

The block diagram of the developed IDSS (DSS, knowledge base) for assessing the technical condition of the CTS is shown in Fig. 1.

When developing a IDSS, a ship's CTS, or more precisely a ship's power plant (SPP), was chosen as the object for assessing the reliability of the technical condition. Such a system is one of the main CTS out of almost hundreds of technical systems installed on the ship.

Assessing the reliability of the SPP needs to take into account the fact that the CTS is characterized by a large number of diagnosed parameters that differ in information content and degree of accessibility, as well as specific and varied operating conditions. In addition, the CTS is characterized by insufficient information about its technical condition.

The functioning of the developed DSS is based on an assessment of the risk of failure of elements and components of the CTS. Those on criteria that reflect taking into account the specifics of the interaction of various elements and components, the correlation of changes in the values of their parameters under various emergency operating conditions of a complex system.

The developed IDSS (Fig. 1) evaluates the reliability of the system using a unified system of parameters of the elements and components of the control system.

IDSS cores are: database; a knowledge base with methods for calculating reliability indicators (probabilities and risks of failures) and a set of decision rules for selecting appropriate decision-making methods; intellectualization model for assessing the technical condition of CTS elements and components.

The basis for constructing an IDSS is the formulation of the decision-making problem in general form:

$$N = f(F, G, A, FR, SG, P, C, PC, NS), \qquad (1)$$

where *F* – many failures of elements and components of the CTS;

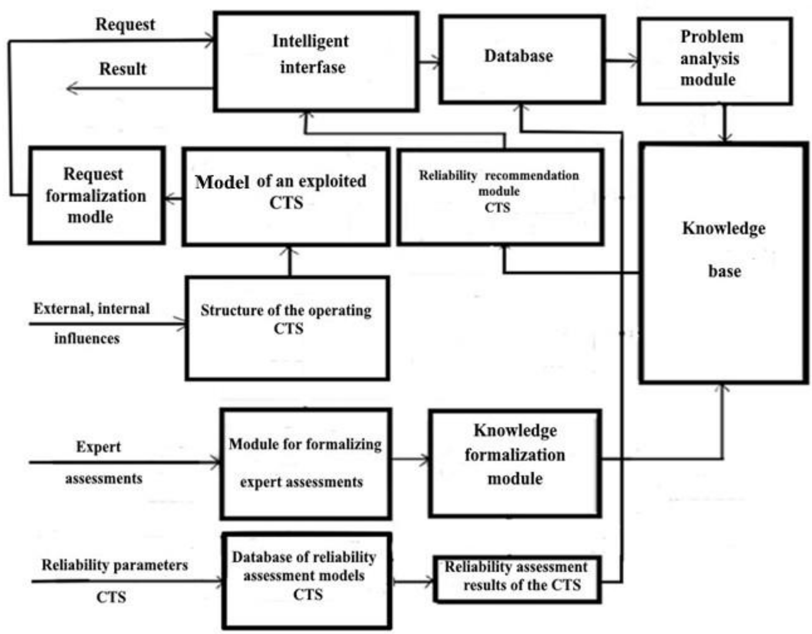


Fig.1. Block diagram of knowledge base, DSS for assessing the technical condition of CTS

G – many set goals (to ensure the reliability of the CTS);

A – many possible alternatives;

FR – multiple failure levels of elements and components of the CTS;

SG, *P*, *C* – set of characteristics, preferences, criteria for ensuring the reliability of elements and components of the CTS;

PC – many principles for coordinating the assessment of alternatives based on individual criteria;

NS – necessary solution to the problem

Preference *F* – assessment of the usefulness of the method of achieving the goal.

The assessment is specified without highlighting the characteristics by which it is made or the characteristics *SG*. The characteristics include the degree of achievement of the goal. To make the final choice of how to achieve the goal, it is necessary to formulate criteria *C*, the number of which is determined by the number of features.

If multiple criteria are used in the IDSS, then it is necessary to apply the principles of *PC* coordination to agree on the assessment of alternatives for each criterion.

To support decision-making on assessments of the risk of CTS failures based on a priori and a posteriori data, as well as when searching for failed elements and system components in order to increase the efficiency of their operation, a method based on dynamic Bayesian trust networks (DBTN) is used [13, 14]. The use of DBTN makes it possible to determine with great accuracy the elements and components of the CTS that are closest to the critical state and their failure.

The task is solved by using a constant system of polling all elements of the system at its various levels for a specific period of time.

This allows, with the help of DBTN, to study extreme situations and accurately determine the critical values of the risk of failure of elements and components of the CTS.

The construction and study of the DBTN probability of loss of performance, assessment of the risk of failure of elements and components of the CTS was carried out using the GiNIe software product [20]. The decision support strategy used when searching for failures of elements and components of ship CTS consists of a number of stages (Fig. 2).

At the initial stage, the numerical values of preliminary assessments of failures of elements and components of the CTS are determined using a diagnostic model based on DBTN. The input variables for the Bayesian diagnostic model are test results.

The model of the operating CTS in the intelligent system for assessing the risk of failure of system components (Fig. 1) in the form of DBTN.

Currently, Bayesian belief networks are actively developing in the field of modeling and knowledge representation [13, 14].

When solving the problems of diagnosing CTS, BBN allow taking into account both the dependence between the types of technical systems and diagnostic features, taking into account the reliability of their checks, and the results of checking diagnostic features, data on failures of CTS components.

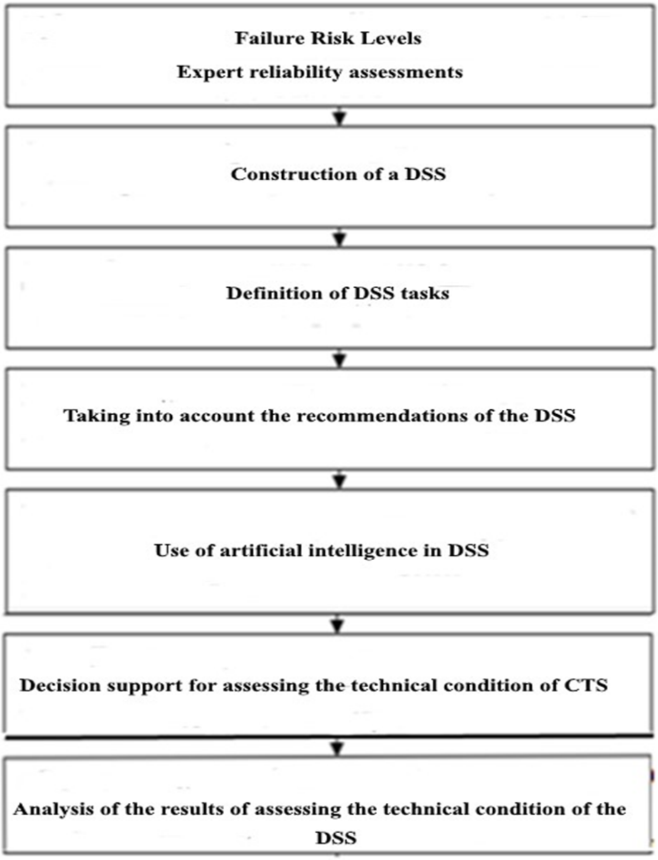


Fig. 2. Strategy for decision support when searching for failures in CTS

The model of an intelligent system for assessing and predicting the risk of failure of components of a complex technical system in the form of a BBN can be written as:

$$< M , S , R , L >, \qquad (2)$$

where *M* - is the set of subsystems (elements) of the CTS;

S - a set of intersystem (interelement) links of CTS;

R - a set of diagnostic assessments of the risk of failures of subsystems (elements), intersystem (interelement) links of CTS;

L - mapping of connections between the sets *M*, *S* and *R*, based on the CTS diagnostic model.

The set of subsystems (elements) of ship CTS, taking into account the hierarchical levels of subsystems (elements), is determined by:

$$M = \{\upsilon_{i_{S(E)}}^{<j_{S(E)}>} \mid i_{S(E)} = \overline{1, I_{S(E)}}; j_{S(E)} = \overline{0, J_{S(E)}}\}, \qquad (3)$$

where $\upsilon_{i_{S(E)}}^{<j_{S(E)}>}$ - is the state of each subsystem (element) of the CTS;

$i_{S(E)}$ - number of subsystem (element) of CTS;

$j_{S(E)}$ - number of the hierarchical level of the subsystem (element) of the CTS;

$I_{S(E)}$ - number of subsystems (elements) of CTS;

$J_{S(E)}$ - number of hierarchical levels of subsystems (elements) of CTS.

The state of each subsystem (element) of the CTS:

$$\upsilon_{i_{S(E)}}^{<j_{S(E)}>} = \{F_{\upsilon_{n_{S(E)}}}, F_{\upsilon_{i_{S(E)}}}, a_{\upsilon_{in_{S(E)}}}, a_{\upsilon_{on_{S(E)}}}\}, \qquad (4)$$

where $F_{\upsilon_{n_{S(E)}}}$ - is the nominal performance of the subsystem (element) of the STS;

$F_{\upsilon_{i_{S(E)}}}$ - operability of a subsystem (element) in case of its partial loss;

$a_{\upsilon_{in_{S(E)}}}, a_{\upsilon_{on_{S(E)}}}$ - intersystem (interelement) connections incoming and outgoing to subsystems (elements), *in, on* – sequence number of incoming and outgoing intersystem (interelement) connections.

A set of intersystem (interelement) links of CTS:

$$S = \{\omega_{c,h}^{<b,q>} \mid c = \overline{1,C}; h = \overline{1,H}; b = \overline{1,B}; q = \overline{1,Q}\}, \tag{5}$$

where $\omega_{c,h}^{<b,q>}$ - is the state of each intersystem (interelement) connection;

c – number of intersystem communication;

h - is the number of the interelement bond;

b - is the number of the hierarchical level of intersystem communication;

q - is the number of the hierarchical level of the interelement connection;

C - is the number of intersystem connections;

H - is the number of interelement bonds;

B - is the number of hierarchical levels of intersystem links;

Q - is the number of hierarchical levels of interelement connections

The state of each intersystem (interelement) connection

$$\omega_{c,h}^{<b,q>} = \{F_{\omega_{cn}}; F_{\omega_{cp}}; F_{\omega_{hn}}; F_{\omega_{hp}}\}, \tag{6}$$

where $F_{\omega_{cn}}$ - is the nominal performance of intersystem connections;

$F_{\omega_{cp}}$ - operability of intersystem communication in case of its partial loss;

$F_{\omega_{hn}}$ - nominal performance of the interelement connection;

$F_{\omega_{hp}}$ - operability of intersystem communication in case of its partial loss.

A set of diagnostic assessments of the risk of failures of subsystems (elements), intersystem (interelement) links of CTS:

$$R < P, Y >$$
$$R_m = \{r_m \mid m = \overline{1,M}\}, \tag{7}$$
$$R_s = \{r_s \mid s = \overline{1,S}\},$$

where *M, S* - are determined based on the failure trees, presented as a set of risk of failures of subsystems (elements) and intersystem (interelement) links, taking into account their failure probabilities (*P*) and damages from failures (*Y*);

r_m - risk of failures of subsystems (elements) of CTS;

r_s - risk of failures of intersystem (interelement) connections.

The initial data for constructing a model of an intelligent system for assessing and predicting the risk of failures of components of a complex technical system on the example of a ship power plant (SPP) [21], based on a dynamic BBN, are: SPP scheme; the principle of operation of the SPP; probability of failures of CTS components.

The construction and study of the BBN of the probability of loss of working capacity, assessments of the risk of failures of CTS components was carried out using the GenNIe software product [20]. It is a fully portable C++ class library that implements graphical decision theory methods such as the Bayesian network.

Jobs and impact diagrams that are directly amenable to inclusion in intelligent systems. Its Windows user interface, Genie is a versatile and user-friendly development environment for graphical decision theory models. modeling tools into intelligent systems.

The use of the GenNIe environment allows diagnosing each component of the CTS.

Perform a regression analysis of the influence of each parent element of the network on its corresponding child element. Implement a graphical display of the results of predicting the risk assessment of failures of CTS components.

Calculate the value of the probability of loss of performance, damage and risk assessments of failures of CTS components.

When modeling the BBN of the SPP (Fig. 3), for various values of the probability (risk) of failure of the input element, the values of the probability (risk) of failures, the performance of the components of the SPP for 20,000 hours of its operation are determined. Symbols of the elements of the SPP are given in Table 1.

Table1

Symbols of the components of the SPP

Component name	Symbol	Failure risk value
Input element	VHOD	0,26
Manual control of the main engine	RUGD	0,035
Compressed air system	SSV	0,047
Control system for propulsion and steering complex (PSC)	SUDRK	0,081
Boiler plant	KU	0,13
Ship power plant	SE	0,09
Fire fighting system	PS	0,01
Main engine	GD	0,16
Remote automated control system of the main engine	DAU	0,01
Ballast drainage system	BOS	0,019
Transfer of power from the main engine to the propeller	PM	0,003
Emergency drive PSC	AP	0,01

The operating state and failure, for example, of the SSV subsystem for the risk of failure at the input element of the SPP 0.014 is shown in Fig.4.

From the retrospective analysis of the research results in the simulation of the SPP, the components that affect the overall performance of the system are identified.

In the study of emergency situations, the analysis of incidents in the CTS, the main goal is to determine the cause of the accident.

It follows from the research results that the maximum non-operating state during the operation of the SPP is 20,000 hours. corresponds to the SUDRK complex (Fig. 3).

Because Since the SUDRK complex is dependent at the level of the hierarchical structure of the SPP, it is necessary to check the complex in order to find the cause of its failure.

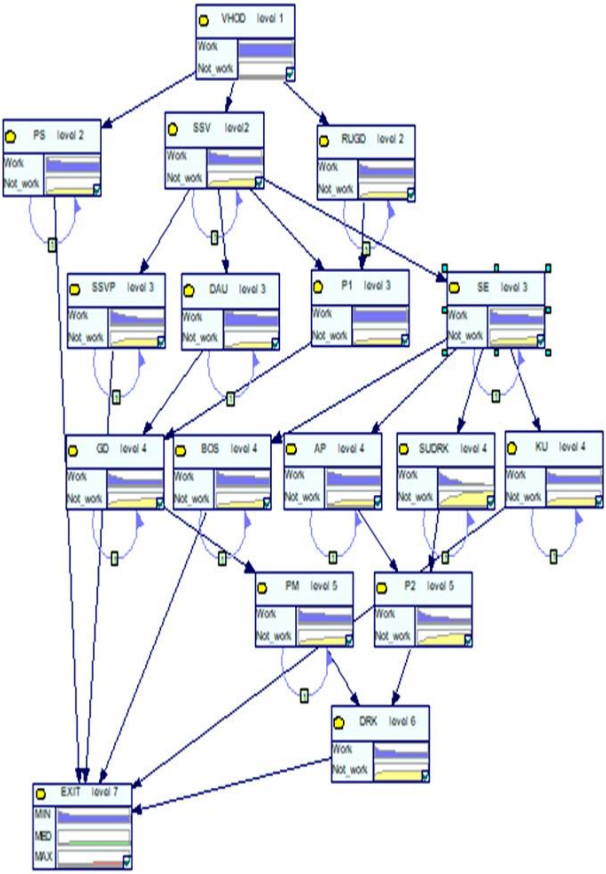


Fig.3. BBN SPP in the GeNIe environment when searching for the causes of failures at the risk of failure of the input component

The purpose of using the BBN in assessing both the probability of loss of performance and the risk of failure of the elements of the CTS components is an a posteriori conclusion.

The a priori data are dynamically recalculated and form a posterior failure risk estimate, which is a priori information, to process the new information.

The posteriori conclusion is based on the procedures for analyzing the data obtained as a result of using the BBN.

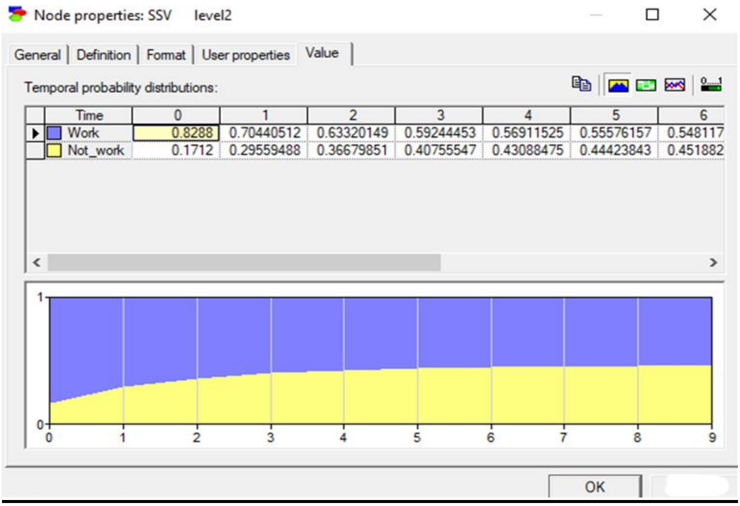


Fig.4. Operating state and failure of the SSV subsystem for the risk of failure at the input element of the SPP 0.014

When implementing this approach in research, modeling using a priori and a posteriori data, the subsystems of the power plant are determined that have the greatest impact on the performance of the main engine and the operation of the entire system for various periods of time.

Figure 5 shows a priori and a posteriori data and studies of the compressed air system for 100 hours of SPP operation. The risk of system failure increased slightly, changing from 0.08 to 0.085.

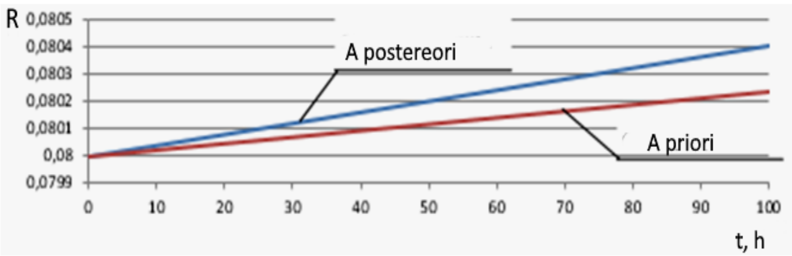


Fig.5. A posteriori and a priori estimates of the risk of failure power plant compressed air systems

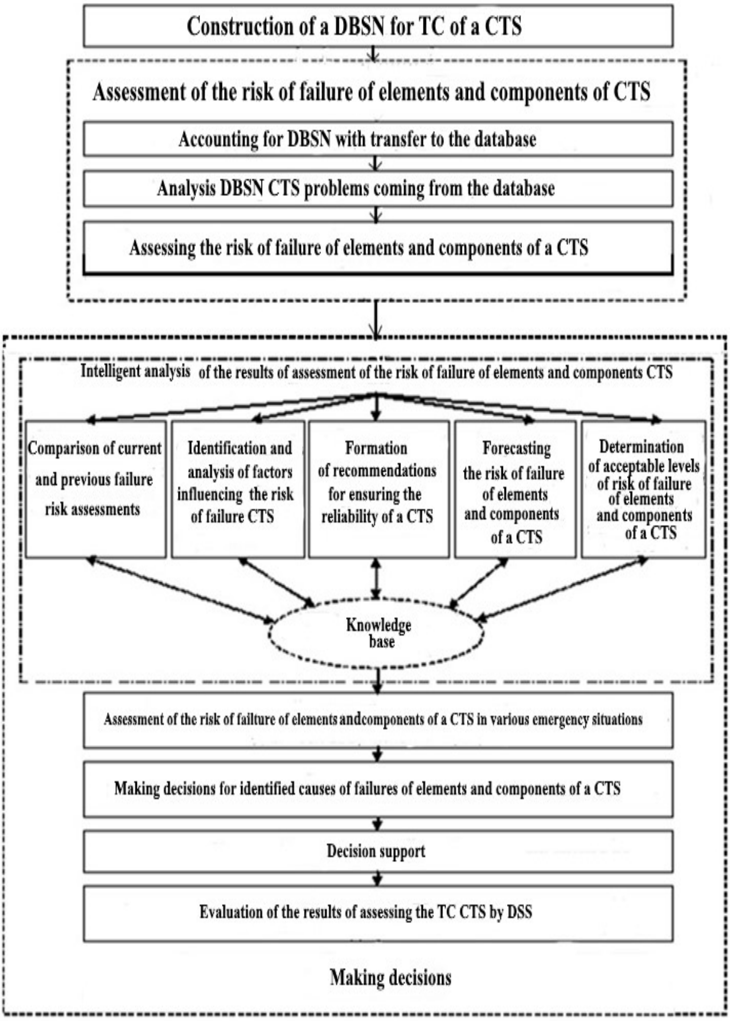


Fig. 6. IDSS algorithm when searching for failures of ship CTS

The implementation of the strategy in the IDSS scheme for assessing the technical condition of the CTS (Fig. 1) is ensured by targeted actions in accordance with the IDSS algorithm (Fig. 6) when searching for failures of elements and components based on assessments of the risk of failure of the diagnosed CTS. As a result of the functioning of the intelligent IDSS for vehicle assessment (using the example of a ship's CTS) in accordance with the algorithm shown in Fig. 6, using the SPP model in an intelligent system (Fig. 1) and DBTN, the dependences of the risk of failure are determined for different samples of failure probabilities of elements and components of systems serving the SPP (Fig. 7, Fig. 8).

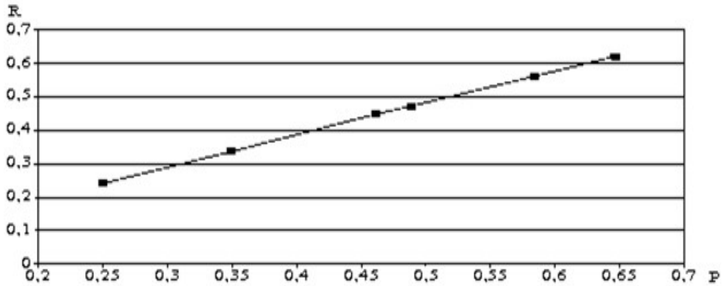


Fig. 7. Dependence of the risk of system failure on the probability of failure of elements of the SPP oil system

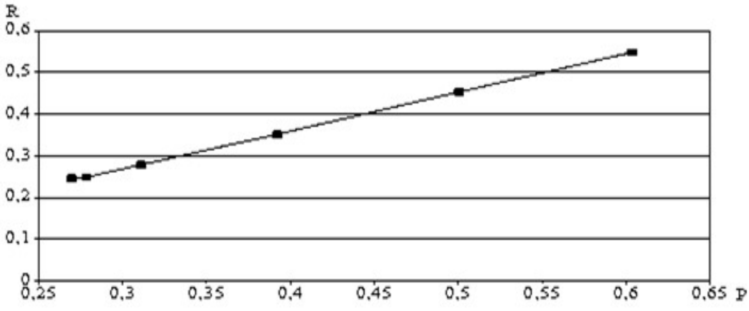


Fig. 8. Dependence of the risk of system failure on the probability of failure of elements of the SPP compressed air system

The problem-oriented knowledge base model is based on the following lists:

– elements. components affecting the trouble-free operation of the CTS;

– states in which the CTS may be in the process of failure-free operation of elements and system components;

– factors under the influence of which the current reliability of the CTS may change, systems transition to a state of failure with disruption of reliable operation;

– problem states into which the CTS can go under the influence of failures of elements and components.

The knowledge base can be presented in the form of a five-level hierarchical tree (Fig. 9).

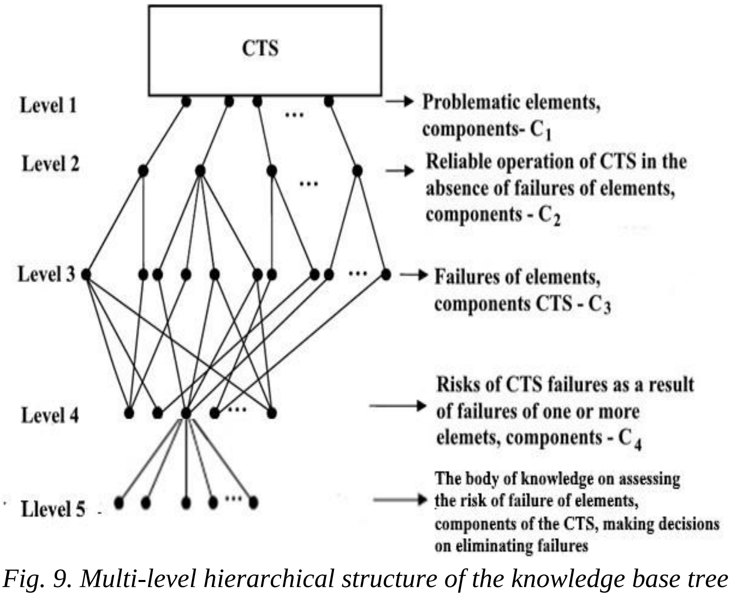


Fig. 9. Multi-level hierarchical structure of the knowledge base tree
Taking into account the hierarchical structure of the knowledge base allows you to quickly localize the cause of a defect or failure and reduce the time for diagnosing CTS.

The acquisition and addition of knowledge is carried out automatically during training and implementation of the expert system.

Filling with knowledge is provided by an expert, as well as by adapting the knowledge base to changes in the subject area and the conditions of its functioning. This is implemented by replacing rules or information in the knowledge base of the IDSS.

5. Conclusions

The proposed decision support system contains: a knowledge base with methods for calculating reliability indicators (probabilities and risk of failures); results of determining the probabilities and risk of failures of elements and components of complex technical systems; intellectualization model for assessing the technical condition of elements and components.

The proposed algorithm for the functioning of a decision support system implements the task of automating the process of assessing the technical condition of complex systems. The use of an intelligent decision support system for assessing the technical condition of complex systems makes it possible to establish the degree of risk of failure of elements and components of the CTS, which increases the efficiency of the systems.

The use of the proposed decision support system for assessing the technical condition of complex systems will improve the reliability of operating systems with insufficient information about their technical condition. Application of the research results of the developed model for the purpose of a retrospective analysis of emergency situations at CTS makes it possible to improve the reliability of systems operation by solving the problem of determining their causes. The application of the developed model, taking into account the hierarchical levels of subsystems (components), intersystem (interelement) connections for an intelligent system for assessing and predicting the risk of failures of components of a complex technical system when searching for the causes of failures of CTS components, allows control the values of the risk of failures of the system components upon receipt of information about failures.

6. References

[1] ISO 13381-1:2015 Condition monitoring and diagnostics of machines – Prognostics – Part 1: General guidelines. – Enter. 2015-03-01, 2015.
[2] V.V. Vychuzhanin, N.D. Rudnichenko, Metody informatsionnykh tekhnologiy v diagnostike sostoyaniya slozhnykh tekhnicheskikh sistem. Monografiya. Odessa: Ekologiya, 2019.

[3] V.V. Vyuzhuzhanin, N.D. Rudnichenko, N.O. Shibaeva, Data Control in the Diagnostics and Forecasting the State of Complex Technical Systems. Herald of Advanced Information Technology 2 2 (2019) 183-196.
[4] V. Vyuzhuzhanin, I. Gritsuk, The Complex Application of Monitoring and Express Diagnosing for Searching Failures on Common Rail System Units DP. SAE International. 2018. 9.10. 2018-01-1773.
[5] V. Vychuzhanin, N. Rudnichenko, Complex Technical System Condition Diagnostics and Prediction Computerization. CMIS-2020 Computer Modeling and Intelligent Systems/ Ceur-ws.org. 2608.2020, pp.1-15.
[6] B.A. Andersen, Diagnostic System for Remote Real-Time Monitoring of Marine Diesel-Electric Propulsion Systems. Leipzig, 2011.
[7] J. Sшrensen Asgeir, Marine Control Systems Propulsion and Motion Control of Ships and Ocean Structures. Department of Marine Technology, NTNU, 2013.
[8] P. Zhang, Z. GaoCao, L. Dong, Marine Systems and Equipment Prognostics and Health Management. Systematic Review from Health Condition Monitoring to Maintenance Strategy. Machines 10 (2022) 72. DOI:10.3390/machines10020072.
[9] V. Vychuzhanin, N. Rudnichenko, Devising a method for the estimation and prediction of technical condition of ship complex systems. Eastern-European Journal of Enterprise Technologies 84 (2016) 4-11.
[10] I. Lazakis, Advanced ship systems condition monitoring for enhanced inspection, maintenance and decision making in ship operations. Transportation Research Procedia 14 (2016) 1679 – 1688.
[11] V. Vychuzhanin, N. Rudnichenko, Assessment of risks structurally and functionally complex technical systems. Eastern-European Journal of Enterprise Technologies 1 2 (2014) 18-22. DOI: 10.15587/1729-4061.2014.19846.
[12] M. Zhang, J. Montewka, T. Manderbacka, P. Kujala, S. Hirdaris, A Big Data Analytics Method for the Evaluation of Ship - Ship Collision Risk reflecting Hydrometeorological Conditions. Reliability Engineering & System Safety 213 (2021). DOI: https://doi.org/10.1016/j.ress.2021.107674.
[13] F.V. Jensen Bayesian, Networks and Decision Graphs. Berlin: Springer, 2007.
[14] R.W. Chen, C. Guan, A Bayesian inference-based approach for performance prognostics towards uncertainty quantification and its applications on the marine diesel engine. ISA Trans. (2021) 159–173.
[15] F. Lan, Y. Jiang, Performance Prediction Method of Prognostics and Health Management of Marine Diesel Engine, in: Proceedings of the 2020 3rd International Conference on Applied Mathematics, Modeling and Simulation, Shanghai, China, 20–21 September 2020/ IOP Publishing: Bristol, UK, 2020.
[16] Y. Chenguang, N. Jing, Neural Network for Complex Systems: Theory and Applications. 2018 | Article ID 3141805. DOI:10.1155/2018/3141805.
[17] G. Phillips-Wren, Intelligent Decision Support Systems. Innovations in Knowledge Management. 2018. DOI:10.1002/9781118522516.ch2
[18] L. Adelman, Evaluating Decision Support and Expert Systems. John Wiley & Sons, Hoboken, NJ. 1992.
[19] A. Ahmad, O. Basir, K. Hassanein, Intelligent expert systems approach to layout decision analysis and design under uncertainty. Intelligent Decision Making: An AI-Based Approach. Springer Berlin, 2008.
[20] Genie. [Online] https://www.zoolz.com/genie9/.
[21] D. Boullosa-Falces, J.L.L. Barrena, Monitoring of fuel oil process of marine diesel engine. Appl. Therm. Eng. 2017.

УДК 004:37:001:62

ІНТЕЛЕКТУАЛЬНИЙ МОНІТОРИНГ ТЕХНІЧНОГО СТАНУ СКЛАДНИХ СИСТЕМ

Вичужанін А.
Національний університет "Одеська політехніка"
E-mail:: v.v.vychuzhanin@op.edu.ua

Анотація. *Актуальність теми обумовлена необхідністю прийняття рішень щодо забезпечення надійності елементів і вузлів складних технічних систем при недостатній інформації про їх технічний стан. Ефективність вирішення проблеми полягає у використанні інформаційних технологій та методів штучного інтелекту, зокрема експертних систем. У статті представлено інтелектуальну систему, яка функціонує з використанням розробленої моделі оцінки та прогнозування ризику відмови компонентів складної технічної системи на прикладі енергетичної установки судна. Побудова моделі з урахуванням ієрархічних рівнів підсистем (компонентів), міжсистемних (міжелементних) зв'язків інтелектуальної системи базується на використанні апріорної інформації про відмови компонентів складних технічних систем. Модель пов'язує види технічного стану компонентів і діагностичні ознаки систем у вигляді ризику їх відмов. Використання апостеріорного висновку в байєсівських мережах переконань дозволяє визначати ризик відмови компонентів системи з урахуванням вхідної діагностичної інформації та інформації про відмови компонентів. Розроблено інтелектуальну систему підтримки прийняття рішень, яка дозволяє оцінювати ризик відмови елементів і компонентів складних технічних систем з використанням елементів штучного інтелекту. Пропонована система підтримки прийняття рішень містить: базу даних; база знань з методами розрахунку показників надійності (ймовірності та ризиків відмов) і набір правил прийняття рішень для вибору відповідних методів прийняття рішень; результати визначення ймовірностей і ризиків відмов елементів і вузлів складних технічних систем з їх ранжуванням; модель інтелектуалізації для оцінки технічного стану елементів і вузлів. Запропонований алгоритм функціонування системи підтримки прийняття рішень реалізує завдання автоматизації процесу оцінки технічного стану складних систем. Результати досліджень моделі оцінки та прогнозування ризику відмови компонентів складної технічної системи підтвердили можливість прогнозування ризику відмови компонентів і системи в цілому. Використання запропонованої системи підтримки прийняття рішень для оцінки технічного стану складних систем дозволить підвищити надійність технічних систем з недостатньою інформацією про їх технічний стан.*

Ключові слова: *складні технічні системи, надійність, ризик відмови, діагностика, підтримка прийняття рішень, прогнозування, інтелектуальна система, байєсовська мережа переконань.*

Section 6. Mathematical and simulation modeling

UDC 511.334:511-33

METHOD FOR COMPUTING EXPONENTIATION MODULO THE POSITIVE AND NEGATIVE INTEGERS

Dr.Sci. Krasnobayev V.A.[1], Ph.D. Yanko A.S[2], Ph.D. Martynenko A.M.[3], Kovalchuk D.M.[4]

[1,4]V. N. Karazin Kharkiv National University, Kharkiv, Ukraine
[2,3]National University «Yuri Kondratyuk Poltava Polytechnic», Poltava, Ukraine
ORCID: 0000-0001-5192-9918, 0000-0003-2876-9316,
0000-0002-9576-0138, 0000-0002-8229-836X
E-mail: [1]v.a.krasnobaev@gmail.com, [2]al9_yanko@ukr.net,
[3]martynenko@pntu.edu.ua [4]kovalchuk.d.n@ukr.net

Abstract. *Numerous publications in recent years indicate that the prospects for creating high-speed computer systems (CS) based on the use of a modular number system (MNS) open up wide opportunities for using CSs with a high degree of parallelization of the processing of integer data. The application of the main properties of MNS and the possibility of using the tabular principle of data processing significantly increases the speed of performing integer arithmetic operations in comparison with the traditional binary positional number system. Research in this area shows the effectiveness of using MNS to increase the speed of execution besides to the arithmetic operations of addition, subtraction and multiplication of integers and the operation of exponentiating of the integers. However, until now there are no effective methods for exponentiating numbers in the MNS in all numeric domain (positive and negative). Therefore, this research developed a system of mathematical ratios that describe the researched process, based on which a method for exponentiating numbers is developed, which unlike the known ones, can be implemented in the negative numeric domain. On the basis of the obtained method, algorithms for exponentiating numbers in the MNS were obtained, according to which devices for their implementation were synthesized.*

Keywords: *arithmetic operation, computer system, system of mathematical ratios, method for exponentiating numbers, modular number system, modular structure, positional number system, tabular multiplication code.*

1. Introduction

Various fields of science and technology require solving computational problems with high accuracy. It is known that rounding errors occur during calculations and its influence increases with the dimension of the problem. Modern CS provides various technical solutions to reduce its impact. However, it is impossible to completely eliminate them if you remain within the framework of the positional number system (PNS). The research of non-positional number systems (modular number system) made it possible to identify a new scientific direction – error-free calculations [1]. Error-free calculations in the MNS are based on representing numbers in the form of residues when dividing them into given prime numbers-modules of the system and performing integer arithmetic operations on it. The final result of calculations in the MNS is converted into the PNS and is presented as a pair of numbers – the numerator and denominator of the irreducible Farey fraction [2]. The advantage of error-free calculations in the MNS is the ability to parallelly process numbers in several modules. This provides a significant increase in performance compared to sequential processing [3]. The effectiveness of multi-module error-free calculations has been proven by many studies, which lies in the use of MNS as a number system of CS for processing integer data [3, 4, 5, 6]. That is, in CS intended for the implementation of integer arithmetic operations of adding, subtracting and multiplying numbers in the positive numerical range, when using MNS, the speed and reliability of solving problems significantly increases [7]. There are many tasks in the CS, the implementation of which requires fast, reliable and high-precision integer arithmetic calculations. However, there is also a large class of tasks and algorithms where, in addition to performing the above integer arithmetic operations in the positive numeric domain, there is a need to implement these operations in the negative numeric domain. One of these operations is the operation of exponentiating the residues of integers to an arbitrary power of a natural number modulo. The lack of an effective method of exponentiating numbers by an arbitrary modulo MNS over the entire numeric domain limits the scope of application of the MNS and therefore computer components in the MNS are rarely implemented in general-purpose computer units. The operation of exponentiating integers a relevant and important operation in such areas of computer science as:

– Mathematical computing, for example, Fourier transform problems and its applications, combinatorial interpretation, etc. [8].

– Modeling of physical processes, for example, when modeling the physical laws of motion.

– Graphics and computer vision: in image processing algorithms.

– Networks and telecommunications, for example, in routing algorithms.

– Machine learning and artificial intelligence: for calculating activation functions in neural networks or data analysis algorithms.

– Geographic information systems: in algorithms, for example in algorithms for calculating the distance between points on a geographic map.

– Computer games: in game mechanics algorithms.

– Financial computing, for example to calculate the future value of investments.

– Cryptography.

The operation of exponentiating integers modulo has the widest practical application in cryptography. Below are some examples of specific applications of this operation:

– RSA encryption: RSA encryption uses modulo exponentiation to secure messages. In this system, each user has a public and a private key. When sending a message, the user encrypts it with the recipient's public key using the modulo exponentiation operation. The recipient, knowing his secret key, decrypts the message, also using the exponentiation modulo operation.

– Diffie-Hellman algorithm: the operation of exponentiating integers is used to create a shared secret key between two participants. Participants choose a random number and the exponentiating this number is performed using the public key of another participant, after which the resulting number is transmitted to another participant. Then other participant also raises the resulting number to the power of its private key to get the shared secret key.

– ElGamal algorithm: operation of exponentiating used to encrypt and decrypt messages. When encryption, the sender randomly chooses a number and raises the recipient's public key to the power of that number, and also raises the plaintext of the message to the power of the sender's private key. The results of the operations are multiplied to get the ciphertext. When decrypting, the recipient raises the first part of the ciphertext to the power of their private key, then uses that number to divide the second part of the ciphertext to get the original plaintext.

– Hash functions: Hash functions are widely used in computer science for data protection and integrity checking. One of the more popular hash functions, SHA-256, uses modulo exponentiation to compute a hash value.

– The operation of exponentiation modulo is also widely used in banking. For example, SSL uses RSA to encrypt data between a browser and a server.

– In addition, the operation of exponentiating integers modulo can be used in data networks to speed up calculations and reduce the amount of transmitted information.

– Pseudo-random number generation: Pseudo-random number generators often use exponentiation modulo to generate a sequence of random numbers. One of the most popular algorithms for generating pseudo-random numbers is the Mersenne algorithm, which uses the operation of exponentiation modulo [8, 9, 10]. In general, the operation of exponentiating integers modulo is one of the key elements of many cryptographic algorithms and has many practical applications in computer science and other fields. At the moment, many modern programming languages do not have tools (operators) that can implement the operation of exponentiating numbers, it is especially difficult to implement this operation for negative numbers [9]. Thus, it was noted that there is a numerous class of tasks and algorithms where, in addition to performing integer basic arithmetic operations and the operation of exponentiating integers modulo in a positive numeric domain, there is a need to implement the above operations in a negative numeric domain. The absence of methods for exponentiating numbers represented in the MNS, both in positive and negative numerical areas, significantly narrows the area of effective use of the MNS as a number system of the CS [3, 6]. So, research is aimed at developing a method for exponentiating numbers modulo MNS are relevant and important. However, practical methods cannot be used to perform the exponentiating operation in the negative numeric domain [11].

2. Recent research and literatures review

Research conducted in recent decades in the field of development of an effective number system at the level of the arithmetic-logical device of the CS by a number of authors (Valah M., Svoboda A., Sabo N., Akushsky I. Ya., Yudytskyi D. I., Nikolaychuk Y. M., Dolgov O. I., Torgashov V. A., Amerbaev V. M., Paulier P., Thornton M.A., Dreschler R., Miller D.M., Hung, C.Y., Parhami, B, etc.) proved that the application of the MNS as a

CS number system showed that the use of the MNS as a CS number system to perform basic arithmetic operations (addition, subtraction and multiplication between integers and real numbers) significantly increases the speed of implementation of the above operations. Research in the field of the application of the non-positional number system (NPNS), to which the MNS belongs, shows that its practical use in the CS allows to significantly increase the productivity of the realization of arithmetic operations. It should be noted that there is a class of problems and algorithms where in addition to performing an integer arithmetic operation (addition, subtraction, multiplication), there is a need to implement the operation of exponentiating numbers in all numeric domain. This problem is discussed in the works of scientists Dr. Dimitrios Schinianakis and Thanos Stouraitis, who explored MNS in cryptography. In particular, in work [12] providing a detailed explanation of cryptographic algorithms based on the MNS, including RSA, ElGamal, and elliptic curve cryptography. Authors discuss the problem and limitation of using the MNS in cryptography, such as the difficulty of handling negative numbers. Various methods have been explored to solve this problem and ensure the reliability of cryptographic systems based on the system of residual classes:

– In article [13] presents the implementation of the ELGamal cryptoalgorithm for information flows encryption/decryption, which is based on the application of the vector-modular method of modular exponentiation and multiplication. This allows us to replace the complex operation of the modular exponentiation with multiplication and the last one with addition that increases the speed of the cryptosystem. In accordance with this, the application of the vector-modular method allows us to reduce the modular exponentiation and multiplication temporal complexity in comparison with the classical one. Despite providing an excellent foundation, the authors do not explore the use of NPNS.

– In article [14], presents a randomized Montgomery Powering Ladder Modular Exponentiation (RMPLME) scheme for side channel attacks (SCA) resistant Rivest-Shamir-Adleman (RSA) and its leakage resilience analysis. This method randomizes the computation time of square-and-multiply operations for each exponent bit of the Montgomery Powering Ladder

(MPL) based RSA exponentiation using various radices (Radix – 2, 2^2, and 2^4) based Montgomery Modular multipliers (MMM) randomly. The randomized computations of RMPLME generates non-uniform timing channels information and power traces thus protecting against SCA. In this work, The authors have developed and implemented a) an unmasked right-to-left Montgomery Modular Exponentiation (R-L MME), b) MPL exponentiation and c) the proposed RMPLME schemes for RSA decryption. All the three realizations have been assessed for side channel leakage using Welch's t-test and analyzed for secured realizations based on degree of side channel information leakage. RMPLME scheme shows the least side-channel leakage and resilient against SPA, DPA, C-Safe Error, CPA and Timing Attacks. Despite providing an excellent foundation, the authors do not explore the use of NPNS.

– In article [15], the authors shows that modular exponentiation (ME) is a complex operation for several public-key cryptosystems (PKCs). Moreover, ME is expensive for resource-constrained devices in terms of computation time and energy consumption, especially when the exponent is large. ME is defined as the task of raising an integer x to power k and reducing the result modulo some integer n. Several methods to calculate ME have been proposed. In this paper, the authors present the efficient ME methods. The authors then implement the methods using different security levels of RSA keys on a Raspberry Pi. Finally, authors give the fast ME method. But the method proposed by the authors is not possible exponentiation modulo the negative integers.

– In article [16], the authors present a reconfigurable architecture for pre-computation methods to compute modular exponentiation and thereby speeding up RSA and Diffie-Hellman like protocols. The authors choose Diffie-Hellman key pair $(a, g^a \bmod p)$ to illustrate the efficiency of Boyko et al's scheme in hardware architecture that stores pre-computed values a_i and corresponding g_i^a in individual block RAM. The authors use a Pseudo-random number generator (PRNG) to randomly choose a_i values that are added and corresponding g_i^a values are multiplied using modular multiplier to arrive at a new pair $(a, g^a \bmod p)$. Further, authors present the advantage of using Montgomery and interleaved methods for batch multiplication to optimise time and area. The authors show that a 1024-bit modular exponentiation can be performed in less than 73μs at a clock rate of 200MHz on a Xilinx Virtex 7 FPGA. But the method proposed by the authors is not possible exponentiation modulo the negative integers.

– In article [17], the authors show that modular exponentiation is one of the most fundamental operations in lots of encryption and signature systems. Due to the heavy computation cost of modular exponentiation, many schemes have been put forward to securely outsource modular exponentiation to cloud. However, most of the existing approaches need two non-colluded cloud servers to securely complete the modular exponentiation, which will result in private data leakage upon the cloud servers collude. Besides, most existing schemes assume both base and exponent in modular exponentiation are private, which does not conform to many real-world applications. For example, in public key encryption system that uses modular exponentiation, one of base and exponent is the public key, and the other is a message. Usually, only the message should be privately protected. In this paper, the authors propose two secure outsourcing schemes for fixed-base (public base and private exponent) and fixed-exponent (private base and public exponent), respectively. In the proposed schemes, the authors employ only one cloud server and can thus avoid collusion attack. Further, the authors achieve an efficient and secure Paillier encryption outsourcing scheme based on our secure modular exponentiation outsourcing methods. Additionally, the authors theoretically analyze our overheads and leverage simulation experiments to evaluate our proposed solutions, which show our schemes can achieve high efficiency. But the method proposed by the authors is not possible exponentiation modulo the negative integers.

– In article [18], the authors provide a method for providing efficient final exponentiation algorithms for a specific cyclotomic family of curves with arbitrary prime k of $k \equiv 1 \pmod 6$. Applying the proposed method for several curves such as k=7, 13 and 19, it is found that the proposed method gives rise to the same algorithms as the previous state-of-the-art ones by the lattice-based method. But the method proposed by the authors is not possible exponentiation modulo the negative integers.

Indeed, as a review of the literature showed, today the question of searching for alternative methods and algorithms for implementing the operation of exponentiating numbers (integers) is quite popular. A wide arsenal of publications related to this problem is proof of the unresolved choice of the optimal number system, on the basis of which it is possible to build fast and reliable methods for exponentiating numbers in all numeric domain.

3. Solving the research problem

Thus, the unsolvedness of the above-listed problem, connected primarily with the operation of exponentiating numbers in the negative numeric domain in MNS, the purpose and the objectives of the research.

The purpose of the research to develop an effective system of mathematical ratios (SMR) of the process of exponentiating numbers in the MNS in all numeric domain.

The main objective of the research is to develop a method for exponentiating the residues of integers by an arbitrary modulo MNS to a power of a natural number, which, unlike the known ones, is based on the properties of the NPNS, by using the tabular principle of data processing, which increases the performance (speed) of the operation of exponentiating integers to a power.

The object of the research is data processing processes in the MNS, presented in integer form.

The subject of the research is methods for exponentiating numbers (integers) by an arbitrary modulo MNS.

This paper uses *research methods* related to number theory, the principles of system analysis, the theory of computational processes and systems, the results of the Chinese remainder theorem. The research is based on the application of the properties (independence and their low bit-depth of the residues) of the MNS. Let's briefly consider the influence of each of the properties of the MNS that were used. The independence of the residues makes it possible to construct a CS in the form of a set of independent, parallel working separate computational information processing paths, operating independently of each other according to their specific module. Thus, the CS operating in the MNS has a modular design, which allows for maintenance and elimination of failures and malfunctions of the computing paths by simply replacing them without interrupting the solution of the computing problem. The time of implementation of arithmetic operations in the CS is determined by the time of implementation of the operation in the computing path according to the greatest base (module) of the MNS. In result in the MNS, each number appears as several low-order positional numbers, which are the residues from dividing the original number by mutually prime bases. Unlike the usual binary PNS, where basic arithmetic operations are carried out sequentially by digits, beginning with the least significant, the MNS permits for the parallelization of this process. All operations over the residues for each base are conducted separately and independently, making it easy and fast due to their low bit-depth. The low bit-depth (low-bit capacity) of the residues allows to significantly increase the reliability of the CS and the speed of performing arithmetic operations, both due to the low bit-depth of the computing paths of the CS, and due to the possibility of using (unlike PNS) tabular (table-based) arithmetic, where arithmetic operations of addition, subtraction and multiplication are performed almost in one machine cycle. The peculiarity of tabular arithmetic is that the result of the operation is not calculated every time. Instead, it's calculated once, stored in a memory device, and then retrieved as needed. Therefore, an operation in RNS with tabular arithmetic is carried out within one period of the synchronizing frequency (machine cycle). In particular, the low bit-depth of the residues in the representation of numbers in modular arithmetic allows for a wide choice of options for system-technical solutions when implementing modular arithmetic operations based on the following principles: Adder principle (based on low-bit binary adders). Tabular principle (based on the use of small-sized Read-Only Memory (ROM) units). Ring shift principle (based on the use of ring shift registers).

The tabular principle in MNS allows not only the implementation of basic operations but also complex functions, all within one machine cycle.

This prompts one of the incredible properties of MNS: the effective performance of a modular CS can be significantly, even exponentially, higher than a positional system with the same clock frequency. Indeed, an operation that a typical binary positional system performs in 100 cycles, system in the MNS executes in a single cycle, naturally resulting in a 100-fold increase in effective performance under otherwise equal conditions [7].

There are other important properties of the MNS that determine the positive and unique features (for example, the ability to have different reliability, accuracy of calculations and speed in the dynamics of the computing process) of modular arithmetic, but in the framework of this research, it was the independence and low bit-depth of the residues that were used. Based on an analysis of the possible use of the listed properties, MNS has significant advantages over PNS. One of which is the lower computational and time complexity of integer arithmetic operations, including the operation of exponentiation, which consists of multiplication operations.

The principle of realization of the system of mathematical ratios is developed using the above-considered properties of the MNS that define the non-positional code data structure of the MNS which ensures high productivity (speed) and reliability of arithmetic operations for the implementation of computational algorithms in the CS, consisting of a set of arithmetic (modular) operations [19].

A distinctive feature of this research is that the proposed system of mathematical ratios of the process of exponentiating numbers in all numeric domain will significantly expand the scope of MNS and increase the speed of this operation in CS, and the method are reduced to algorithms, on the basis of which classes of patent-eligible devices for which Ukrainian patents have been obtained have been developed [20, 21].

Solving the problem is to use the proposed principles and properties of MNS, on the basis of which the method for exponentiating numbers in the MNS in all numeric domain is implemented. In this case, the advantage of the proposed method should lie not only in the possibility of implementation in the negative numeric domain, but also in the speed of execution (due to the use of the tabular principle of data processing).

4. Realization of computing exponentiation of numbers in the MNS in all numeric domain

By the type of the original number presented in the MNS $C_{MNS} = (c_1 \| c_2 \| ... \| c_{k-1} \| c_k \| c_{k+1} \| ... \| c_l)$, where $\|$ – mathematical sign of the concatenation operation: gluing operation, joining operation; it is difficult to attribute the number to any of the numeric domains. Consider the options of finding numbers in the MNS in the required numeric domains.

The first option. The original number represented in the MNS C_{MNS} has an additional two (or one) sign bits $\Psi_{+C_{MNS}}$ and $\Psi_{-C_{MNS}}$, where:

$$\Psi_{+C_{MNS}} = \begin{cases} 1, if\ C_{MNS} > 0, \\ 0, if\ C_{MNS} < 0; \end{cases}$$

$$\Psi_{-C_{MNS}} = \begin{cases} 0, if\ C_{MNS} > 0, \\ 1, if\ C_{MNS} < 0. \end{cases}$$

In this case, the original number in the MNS will be represented as $C_{MNS} = [\Psi_{+C_{MNS}} \| \Psi_{-C_{MNS}} \| (c_1 \| c_2 \| ... \| c_{k-1} \| c_k \| c_{k+1} \| ... \| c_l)]$. For this option, it is technically difficult to determine the sign of the result of the operation [5]. The second option. When performing the operation of exponentiating numbers by an arbitrary modulo MNS in all numeric domain, it is supposed to convert the number C_{MNS} in a modular structure (MS) $C^{\rightarrow}_{MNS}$ [3]:

$$\begin{cases} C^{\rightarrow}_{MNS} = \frac{1}{2}D + \left|C_{MNS}\right|, if\ C_{MNS} \geq 0, \\ C^{\rightarrow}_{MNS} = \frac{1}{2}D - \left|C_{MNS}\right|, if\ C_{MNS} < 0, \end{cases}$$

i.e. for positive numbers: $C^{\rightarrow}_{MNS} = \frac{1}{2}D + \left|C_{MNS}\right|$ and for negative: $C^{\rightarrow}_{MNS} = \frac{1}{2}D - \left|C_{MNS}\right|$, where $D = \prod_{k=1}^{l} p_k$, p_k – an arbitrary MNS module.

A system of mathematical ratios the process of exponentiating numbers was developed based on an analytical ratio $(C^r_{MNS})^{\rightarrow} = f(C^{\rightarrow}_{MNS})$ which defines the relation of the result C^r_{MNS} number exponentiating operations

C_{MNS} in the MNS (to the degree of *r*) presented in the MS, from the number $C_{MNS}^{\rightarrow}$ immediately in the MS [22].

Let's define SMR $(C_{MNS}^{r})^{\rightarrow} = f(C_{MNS}^{\rightarrow})$ for values $r=2$ and $p_1 = 2$. In this case, in the MNS there is the following:

$$D = \prod_{k=1}^{l} p_k = (0 \,||\, 0 \,||\, ... \,||\, 0 \,||\, ... \,||\, 0), \tag{1}$$

$$\frac{1}{2} D = \prod_{k=2}^{l} p_k = (1 \,||\, 0 \,||\, ... \,||\, 0 \,||\, ... \,||\, 0), \tag{2}$$

where *l* – the number of bases of the MNS.

According to the representation of the MS numbers in the MNS, have that:

$$\begin{cases} C_{MNS}^{\rightarrow} = \frac{1}{2} D + C_{MNS}, \\ (C_{MNS}^{r})^{\rightarrow} = \frac{1}{2} D + C_{MNS}^{r}. \end{cases} \tag{3}$$

When changing the domain of numbers C_{MNS} and $C_{MNS}^{\rightarrow}$ from the mathematical expression (3) can be defined as:

$$(C_{MNS}^{r})^{\rightarrow} = \left(\frac{1}{2} D + C_{MNS}^{r}\right) \bmod D. \tag{4}$$

Let's carry out the following numerical transformations:

$$(C_{MNS}^{\rightarrow})^2 = C_{MNS}^{\rightarrow} \cdot C_{MNS}^{\rightarrow} = (\frac{1}{2} D + C_{MNS}) \cdot (\frac{1}{2} D + C_{MNS}) = C_{MNS}^2 + C_{MNS} \cdot D + \frac{1}{2} D \cdot \frac{1}{2} D. \tag{5}$$

In this case, expression (5) will be presented as:

$$(C_{MNS}^{\rightarrow})^2 = C_{MNS}^2 + \frac{1}{2} D. \tag{6}$$

On the other hand, there is the following:

$$\begin{cases} (C_{MNS}^{\rightarrow})^2 = \frac{1}{2} D + C_{MNS}^2, \\ C_{MNS}^2 = (C_{MNS}^{\rightarrow})^2 - \frac{1}{2} D. \end{cases} \tag{7}$$

By replacing the value of C_{MNS}^2 from relation (7) into relation (6) can obtain that:

$$(C_{MNS}^{\rightarrow})^2 = (C_{MNS}^2)^{\rightarrow} - \frac{1}{2}D + \frac{1}{2}D \quad \Rightarrow \quad (C_{MNS}^{\rightarrow})^2 = (C_{MNS}^2)^{\rightarrow}. \qquad (8)$$

Mathematical ratio (8) is the SMR of squaring numbers modulo MNS. It is also possible to get SMR for the general case, when $r > 2$ [23] in the form $(C_{MNS}^r)^{\rightarrow} = (C_{MNS}^{\rightarrow})^r$.

In this case, it is obvious that:

$$(C_{MNS}^r)^{\rightarrow} = (C_{MNS}^{r-1})^{\rightarrow} \cdot C_{MNS}^{\rightarrow}, \qquad (9)$$

Mathematical ratio (9) is a generalized SMR process of exponentiating numbers modulo MNS. For the convenience of applying ratio (9), it is sometimes possible to use the ratio:

$$\begin{gathered}[c_1^r(\mathrm{mod}\ p_1) \,\|\, c_2^r(\mathrm{mod}\ p_2) \,\|\, \ldots \,\|\, c_k^r(\mathrm{mod}\ p_k) \,\|\, \ldots \,\|\, c_l^r(\mathrm{mod}\ p_1)]^{\rightarrow} = \\ = (c_1^{\rightarrow} \,\|\, c_2^{\rightarrow} \,\|\, \ldots \,\|\, c_{k-1}^{\rightarrow} \,\|\, c_k^{\rightarrow} \,\|\, c_{k+1}^{\rightarrow} \,\|\, \ldots \,\|\, c_l^{\rightarrow})^r\end{gathered} \qquad (10)$$

Processed numbers C_{MNS}^r and $(C_{MNS}^{\rightarrow})^r$ are in the range:

$$\begin{cases} -\frac{1}{2}D \le C_{MNS}^r \le \frac{1}{2}(D-1), \\ 0 \le (C_{MNS}^{\rightarrow})^r \le D-1. \end{cases}$$

Based on SMR (9) implementation of the operation of exponentiating numbers, consider the method of tabular (matrix) realization of the operation of exponentiating numbers modulo MNS in all numeric domain [24-27]. With a tabular realization of the operation of exponentiating numbers modulo p_k MNS, the residues of the number $C_{MNS}^{\rightarrow}$ are encoded by the tabular multiplication code (TMC) [22] as follows $c_k^{\rightarrow} = [\mu_{c_k}^{\rightarrow} \,\|\, (c_k^{\rightarrow})^*]$. The sign $\mu_{c_k}^{\rightarrow}$ of the TMC can be represented in the following form.

For p_k – an even number:

$$\mu_{c_k}^{\rightarrow} = \begin{cases} 0, \ if \ 0 \le c_k^{\rightarrow} \le \frac{p_k}{2}, \\ 1, \ if \ \frac{p_k}{2} < c_k^{\rightarrow} \le p_k - 1. \end{cases} \qquad (11)$$

For p_k – an odd number:

$$\mu_{c_k}^{\rightarrow} = \begin{cases} 0, if\ 0 \le c_k^{\rightarrow} \le \dfrac{(p_k - 1)}{2}, \\ 1, if\ \dfrac{(p_k - 1)}{2} < c_k^{\rightarrow} \le p_k - 1. \end{cases} \qquad (12)$$

The numerical part of $(c_k^{\rightarrow})^*$ TMC of the residue $c_k^{\rightarrow}$ is determined as follows. For p_k – an even number:

$$(c_k^{\rightarrow})^* = \begin{cases} c_k^{\rightarrow}, if\ 0 \le c_k^{\rightarrow} \le \dfrac{p_k}{2}; \\ \overline{c_k^{\rightarrow}} = p_k - c_k^{\rightarrow}, if\ \dfrac{p_k}{2} < c_k^{\rightarrow} \le p_k - 1, \end{cases} \qquad (13)$$

wherein $0 \le (c_k^{\rightarrow})^* \le \dfrac{p_k}{2}$.

For p_k – an odd number:

$$(c_k^{\rightarrow})^* = \begin{cases} c_k^{\rightarrow}, if\ 0 \le c_k^{\rightarrow} \le \dfrac{(p_k - 1)}{2}; \\ \overline{c_k^{\rightarrow}} = p_k - c_k^{\rightarrow}, if\ \dfrac{(p_k - 1)}{2} < c_k^{\rightarrow} \le p_k - 1, \end{cases} \qquad (14)$$

wherein $0 \le (c_k^{\rightarrow})^* \le \dfrac{(p_k - 1)}{2}$.

Result $(c_k^{\rightarrow} \cdot c_k^{\rightarrow}) \bmod p_k$ residue multiplication operations $c_k^{\rightarrow}$ on itself modulo p_k submitted to TMC, i.e. in the form $\left\{\mu_k^{\rightarrow} \,\|\, [(c_k^{\rightarrow})^* (c_k^{\rightarrow})^*] \bmod p_k\right\}$. Then the condition performed $(\mu_{c_k}^{\rightarrow} + \mu_{c_k}^{\rightarrow}) = 0 \pmod 2$. In this case, there is the following:

$$(c_k^{\rightarrow} \cdot c_k^{\rightarrow}) \bmod p_k = [(c_k^{\rightarrow})^* \cdot (c_k^{\rightarrow})^*] \bmod p_k, \qquad (15)$$

wherein $0 \le [(c_k^{\rightarrow})^* \cdot (c_k^{\rightarrow})^*] \bmod p_k \le p_k - 1$.

Based on the developed SMR (9) and using the tabular realization of the modular multiplication operation [22, 28-31], the research improves the method of exponentiating numbers modulo MNS in all numeric domain.

The developed method. A method for exponentiating numbers in the MNS in all numeric domain consists of the following stages:

Set the initial numbers for the realization of the method for exponentiating number $C_{MNS}=(c_1\,||\,c_2\,||\,...\,||\,c_{k-1}\,||\,c_k\,||\,c_{k+1}\,||\,...\,||\,c_l)$, an arbitrary modulo $p_k\ (k=\overline{1,l})$ MNS (to the degree of *r*).

Coding of initial numbers C_{MNS} into code words [32] presented in the MS of the form $C_{MNS}^{\rightarrow}$:

$$\begin{cases} C_{MNS}^{\rightarrow}=\frac{1}{2}D+\left|C_{MNS}\right|, if\ C\geq 0, \\ C_{MNS}^{\rightarrow}=\frac{1}{2}D-\left|C_{MNS}\right|, if\ C<0, \end{cases}$$

$$\begin{cases} -\frac{1}{2}D\leq C_{MNS}\leq\frac{1}{2}(D-1), \\ 0\leq C_{MNS}^{\rightarrow}\leq D-1. \end{cases}$$

$$\begin{cases} (C_{MNS}^{r})^{\rightarrow}=\frac{1}{2}D+\left|C_{MNS}^{r}\right|, if\ C_{MNS}^{r}\geq 0, \\ (C_{MNS}^{r})^{\rightarrow}=\frac{1}{2}D-\left|C_{MNS}^{r}\right|, if\ C_{MNS}^{r}<0, \end{cases}$$

$$\begin{cases} -\frac{1}{2}D\leq C_{MNS}^{r}\leq\frac{1}{2}(D-1), \\ 0\leq (C_{MNS}^{\rightarrow})^{r}\leq D-1. \end{cases}$$

Representation of the residues $c_k^{\rightarrow}$ of the number $C_{MNS}^{\rightarrow}=(c_1^{\rightarrow}\,||\,c_2^{\rightarrow}\,||\,...\,||\,c_{k-1}^{\rightarrow}\,||\,c_k^{\rightarrow}\,||\,c_{k+1}^{\rightarrow}\,||\,...\,||\,c_l^{\rightarrow})$ in the MS by modules $p_k\ (k=\overline{1,l})$ based on the apply of TMC $c_k^{\rightarrow}=[\mu_{c_k}^{\rightarrow}\,||\,(c_k^{\rightarrow})^{*}]$, where:

for p_k – an even number:

$$\mu_{c_k}^{\rightarrow}=\begin{cases} 0, if\ 0\leq c_k^{\rightarrow}\leq\frac{p_k}{2}, \\ 1, if\ \frac{p_k}{2}<c_k^{\rightarrow}\leq p_k-1, \end{cases}$$

$$(c_k^{\rightarrow})^* = \begin{cases} c_k^{\rightarrow}, \; if \;\; 0 \le c_k^{\rightarrow} \le \dfrac{p_k}{2}; \\ \overline{c_k^{\rightarrow}} = p_k - c_k^{\rightarrow}, \; if \;\; \dfrac{p_k}{2} < c_k^{\rightarrow} \le p_k - 1, \end{cases}$$

wherein $0 \le (c_k^{\rightarrow})^* \le \dfrac{p_k}{2}$;

for p_k – an odd number:

$$\mu_{c_k}^{\rightarrow} = \begin{cases} 0, \; if \;\; 0 \le c_k^{\rightarrow} \le \dfrac{(p_k - 1)}{2}, \\ 1, \; if \;\; \dfrac{(p_k - 1)}{2} < c_k^{\rightarrow} \le p_k - 1, \end{cases}$$

$$(c_k^{\rightarrow})^* = \begin{cases} c_k^{\rightarrow}, \; if \;\; 0 \le c_k^{\rightarrow} \le \dfrac{(p_k - 1)}{2}; \\ \overline{c_k^{\rightarrow}} = p_k - c_k^{\rightarrow}, \; if \;\; \dfrac{(p_k - 1)}{2} < c_k^{\rightarrow} \le p_k - 1, \end{cases}$$

wherein $0 \le (c_k^{\rightarrow})^* \le \dfrac{(p_k - 1)}{2}$;

Definition of result $(c_k^{\rightarrow})^2 = (c_k^{\rightarrow} \cdot c_k^{\rightarrow}) \bmod p_k$, $(k = \overline{1, l})$ modular multiplication operations in the form $\mu_k^{\rightarrow} \,||\, [(c_k^{\rightarrow}) \cdot (c_k^{\rightarrow})] \bmod p_k$.

Determining the result of an operation $\left[(C_{MNS}^{\rightarrow})^{r-1} \cdot C_{MNS}^{\rightarrow}\right] \bmod D = \Big\{\left[(c_1^{\rightarrow})^{r-1}\right] \bmod p_1 \,||\, \left[(c_2^{\rightarrow})^{r-1}\right] \bmod p_2 \,||\, ... \,||\, \left[(c_2^{\rightarrow})^{r-1}\right] \bmod p_k \,||\, ...$ $... \,||\, \left[(c_l^{\rightarrow})^{r-1}\right] \bmod p_l \Big\} \cdot (c_1^{\rightarrow} \,||\, c_2^{\rightarrow} \,||\, ... \,||\, c_{k-1}^{\rightarrow} \,||\, c_k^{\rightarrow} \,||\, c_{k+1}^{\rightarrow} \,||\, ... \,||\, c_l^{\rightarrow})$ exponentiating numbers c_k of integer $C_{MNS} = (c_1 \,||\, c_2 \,||\, ... \,||\, c_{k-1} \,||\, c_k \,||\, c_{k+1} \,||\, ... \,||\, c_l)$ modulo p_k $(k = \overline{1, l})$ MNS to the degree *r*.

According to the SMR:

$$\left[c_1^r (\bmod p_1) \,||\, ... \,||\, c_k^r (\bmod p_k) \,||\, ... \,||\, c_l^r (\bmod p_l)\right]^{\rightarrow} =$$

$= (c_1^{\rightarrow} \,||\, c_2^{\rightarrow} \,||\, ... \,||\, c_{k-1}^{\rightarrow} \,||\, c_k^{\rightarrow} \,||\, c_{k+1}^{\rightarrow} \,||\, ... \,||\, c_l^{\rightarrow})^r$ the process of exponentiating numbers modulo, the operation is implemented:

$$\left[(C_{MNS}^{\rightarrow})^{r-1}\cdot C_{MNS}^{\rightarrow}\right]\mathrm{mod}\,D=\Big\{\left[(c_1^{\rightarrow})^{r-1}\right]\mathrm{mod}\,p_1\,\|\left[(c_2^{\rightarrow})^{r-1}\right]\mathrm{mod}\,p_2\,\|\ldots$$

$$\ldots\|\left[(c_k^{\rightarrow})^{r-1}\right]\mathrm{mod}\,p_k\,\|\ldots\|\left[(c_l^{\rightarrow})^{r-1}\right]\mathrm{mod}\,m_l\Big\}\cdot(c_1^{\rightarrow}\,\|\,c_2^{\rightarrow}\,\|\ldots\|\,c_{k-1}^{\rightarrow}\,\|\,c_k^{\rightarrow}\,\|\,c_{k+1}^{\rightarrow}\,\|\ldots\|\,c_l^{\rightarrow})$$

exponentiating numbers modulo p_k MNS in all numeric domain.

5. Results

The result of the research is induced in the form of using the developed method for exponentiating numbers for the MNS with bases $p_1=2,\ p_2=5,\ p_3=7,$ wherein $D=70.$ Total volume of positive codewords C_{MNS} in the MNS presented in Table 1, where C_{PNS} – number in the positional number system. Table 2 submitts the original positive numbers C_{MNS} and numbers in MS $C_{MNS}^{\rightarrow}$.

Table 1

The numbers C_{MNS} in the MNS

C_{PNS}	C_{MNS}			C_{PNS}	C_{MNS}		
	$p_1=2$	$p_2=5$	$p_3=7$		$p_1=2$	$p_2=5$	$p_3=7$
0	0	0	0	35	1	0	0
1	1	1	1	36	0	1	1
2	0	2	2	37	1	2	2
3	1	3	3	38	0	3	3
4	0	4	4	39	1	4	4
5	1	0	5	40	0	0	5
6	0	1	6	41	1	1	6
7	1	2	0	42	0	2	0
8	0	3	1	43	1	3	1
9	1	4	2	44	0	4	2
10	0	0	3	45	1	0	3
11	1	1	4	46	0	1	4
12	0	2	5	47	1	2	5
13	1	3	6	48	0	3	6
14	0	4	0	49	1	4	0
15	1	0	1	50	0	0	1
16	0	1	2	51	1	1	2
17	1	2	3	52	0	2	3
18	0	3	4	53	1	3	4
19	1	4	5	54	0	4	5
20	0	0	6	55	1	0	6
21	1	1	0	56	0	1	0
22	0	2	1	57	1	2	1

Table 1 (continues)

23	1	3	2	58	0	3	2
24	0	4	3	59	1	4	3
25	1	0	4	60	0	0	4
26	0	1	5	61	1	1	5
27	1	2	6	62	0	2	6
28	0	3	0	63	1	3	0
29	1	4	1	64	0	4	1
30	0	0	2	65	1	0	2
31	1	1	3	66	0	1	3
32	0	2	4	67	1	2	4
33	1	3	5	68	0	3	5
34	0	4	6	69	1	4	6

Table 2

The numbers $C^{\rightarrow}$ in the MS

C	$C^{\rightarrow}$	C	$C^{\rightarrow}$	C	$C^{\rightarrow}$	C	$C^{\rightarrow}$	C	$C^{\rightarrow}$
-35	0	-21	14	-7	28	7	42	21	56
-34	1	-20	15	-6	29	8	43	22	57
-33	2	-19	16	-5	30	9	44	23	58
-32	3	-18	17	-4	31	10	45	24	59
-31	4	-17	18	-3	32	11	46	25	60
-30	5	-16	19	-2	33	12	47	26	61
-29	6	-15	20	-1	34	13	48	27	62
-28	7	-14	21	0	35	14	49	28	63
-27	8	-13	22	1	36	15	50	29	64
-26	9	-12	23	2	37	16	51	30	65
-25	10	-11	24	3	38	17	52	31	66
-24	11	-10	25	4	39	18	53	32	67
-23	12	-9	26	5	40	19	54	33	68
-22	13	-8	27	6	41	20	55	34	69

The practical use of the developed method for the MNS with bases $p_1 = 2,\ p_2 = 5,\ p_3 = 7,\ D = \prod_{k=1}^{l} p_k = p_1 \cdot p_2 \cdot p_3 = 2 \cdot 5 \cdot 7 = 70$ is presented in the form of examples.

Example 1. Let $C_{MNS} = 2 = (0 \| 2 \| 2)$ and $r = 2.$ Let's define the value of $C_{MNS}^{r} = 2^2.$ Because $C_{MNS} = 2 > 0,$ then can get that $C_{MNS}^{\rightarrow} = \frac{1}{2} D + C_{MNS} = 35 + 2 = 37 = \ = (1 \| 0 \| 0) + (0 \| 2 \| 2) = (1 \| 2 \| 2) = 37.$ Because $r = 2$ as a result of multiplying the value of $C_{MNS}^{\rightarrow} = (1 \| 2 \| 2)$ by itself $1\cdot 1=1(\bmod 2)$, $2\cdot 2=4(\bmod 5)$ and $2\cdot 2=4(\bmod 7)$, then can get that $(C^{\rightarrow}_{MNS})^2=(1\,||\,4\,||\,4)=39.$

Check:

$(C^{\rightarrow}_{MNS})^2=37^2=37\times 37=1369=39(\bmod 70)=(1\,||\,2\,||\,2)\times(1\,||\,2\,||\,2)=$
$=(1\,||\,4\,||\,4)=39.$

$$(C^2_{MNS})^{\rightarrow}=\frac{1}{2}D+C^2_{MNS},\qquad C^2_{MNS}=(C^2_{MNS})^{\rightarrow}-\frac{1}{2}D,\qquad 2^2=39-35=4,$$

$2^2=4.$

Example 2. Let $C_{MNS}=2=(0\,||\,2\,||\,2)$ and $r=3.$ Let's define the value of $C^r_{MNS}=2^3.$ Because $C_{MNS}=2>0,$ then can get that

$$C^{\rightarrow}_{MNS}=\frac{1}{2}D+C_{MNS}=35+2=37==(1\,||\,0\,||\,0)+(0\,||\,2\,||\,2)=(1\,||\,2\,||\,2)=37.$$

The first iteration of the multiplication gives the result: $C^{\rightarrow}_{MNS}\times C^{\rightarrow}_{MNS}=(C^{\rightarrow}_{MNS})^2$, $(C^{\rightarrow}_{MNS})^2=(1\,||\,4\,||\,4)$, because $1\cdot 1=1(\bmod 2)$, $2\cdot 2=4(\bmod 5)$ and $2\cdot 2=4(\bmod 7)$. After the second multiplication (because $r=3$) of the number $C^{\rightarrow}_{MNS}$ $(C^{\rightarrow}_{MNS})^3=C^{\rightarrow}_{MNS}\times C^{\rightarrow}_{MNS}\times C^{\rightarrow}_{MNS}=(C^{\rightarrow}_{MNS})^2\times C^{\rightarrow}_{MNS}$ can get that $(C^{\rightarrow}_{MNS})^3=(1\,||\,4\,||\,4)\times(1\,||\,2\,||\,2)=(1\,||\,3\,||\,1)=43.$

1. Check:

$(C^{\rightarrow}_{MNS})^3=37^3=50653=43(\bmod 70)=(1\,||\,2\,||\,2)\times(1\,||\,2\,||\,2)\times(1\,||\,2\,||\,2)=$
$=(1\,||\,3\,||\,1)=43.$

2. $$C^3_{MNS}=(C^3_{MNS})^{\rightarrow}-\frac{1}{2}D,\quad 2^3=43-35=8,\quad 2^3=8.$$

Example 3. Let $C_{MNS}=-2\;\left[2=(0\,||\,2\,||\,2)\right]$ and $r=2.$ Let's define the value of $C^r_{MNS}=(-2)^2.$ Because $C_{MNS}=-2<0,$ then

$$C^{\rightarrow}_{MNS}=\frac{1}{2}D-C_{MNS}=35-2=33==(1\,||\,0\,||\,0)-(0\,||\,2\,||\,2)=(1\,||\,3\,||\,5)=33.$$

Because $r=2$ there is the following $1\cdot 1=1(\bmod 2)$, $3\cdot 3=4(\bmod 5)$,

$5 \cdot 5 = 4(\bmod 7),$ then can get that $(C^{\rightarrow}_{MNS})^2 = (1 \| 4 \| 4) = 39.$

Check:

$(C^{\rightarrow}_{MNS})^2 = 33^2 = 33 \times 33 = 1089 = 39(\bmod 70) = (1 \| 3 \| 5) \times (1 \| 3 \| 5) =$
$= (1 \| 4 \| 4) = 39.$

$$C^2_{MNS} = (C^2_{MNS})^{\rightarrow} - \frac{1}{2}D,\ (-2)^2 = 39 - 35 = 4,\ (-2)^2 = 4.$$

Example 4. Let $C_{MNS} = -2\ \left[2 = (0 \| 2 \| 2)\right]$ and $r = 3.$ Let's define the value of $C^r_{MNS} = (-2)^3.$ Because $C_{MNS} = -2 < 0,$ then there is the following

$$C^{\rightarrow}_{MNS} = \frac{1}{2}D - C_{MNS} = 35 - 2 = 33 = (1 \| 0 \| 0) - (0 \| 2 \| 2) = (1 \| 3 \| 5) = 33.$$

The first iteration of the multiplication gives the result: $C^{\rightarrow}_{MNS} \times C^{\rightarrow}_{MNS} = (C^{\rightarrow}_{MNS})^2,$ $1 \cdot 1 = 1(\bmod 2),$ $3 \cdot 3 = 4(\bmod 5),$ $5 \cdot 5 = 4(\bmod 7).$ The second iteration (because $r = 3$): $(C^{\rightarrow}_{MNS})^2 \times C^{\rightarrow}_{MNS}.$ In this case, there is the following $(C^{\rightarrow}_{MNS})^2 = (1 \| 4 \| 4),$ $(C^{\rightarrow}_{MNS})^3 = (C^{\rightarrow}_{MNS})^2 \times C^{\rightarrow}_{MNS} =$ $= (1 \| 4 \| 4) \times (1 \| 3 \| 5) = (1 \| 2 \| 6) = 27.$

Check:

$(C^{\rightarrow}_{MNS})^3 = 33^3 = 35937 = 27(\bmod 70) = C^{\rightarrow}_{MNS} \times C^{\rightarrow}_{MNS} \times C^{\rightarrow}_{MNS} = (1 \| 3 \| 5) \times$
$\times (1 \| 3 \| 5) \times (1 \| 3 \| 5) = (1 \| 2 \| 6) = 27.$

$$C^3_{MNS} = (C^3_{MNS})^{\rightarrow} - \frac{1}{2}D,\ (-2)^3 = 27 - 35,\ (-2^3) = -8.$$

Example 5. Let $C_{MNS} = -3\ \left[3 = (1 \| 3 \| 3)\right]$ and $r = 3.$ Let's define the value of $C^r_{MNS} = (-3)^3.$ Because $C_{MNS} = -3 < 0,$ then

$$C^{\rightarrow}_{MNS} = \frac{1}{2}D - C_{MNS} = 35 - 3 = 32 = \quad = (1 \| 0 \| 0) - (1 \| 3 \| 3) = (0 \| 2 \| 4) = 32.$$

The first iteration of the multiplication gives the result: $C^{\rightarrow}_{MNS} \times C^{\rightarrow}_{MNS} = (C^{\rightarrow}_{MNS})^2 = (0 \| 2 \| 4) \times (0 \| 2 \| 4) = (0 \| 4 \| 2).$ The second iteration (because $r = 3$): $(C^{\rightarrow}_{MNS})^2 \times C^{\rightarrow}_{MNS} = (C^{\rightarrow}_{MNS})^3 = (0 \| 4 \| 2) \times (0 \| 2 \| 4) = (0 \| 3 \| 1).$ Thus, can get that $C^{r}_{MNS} = (-3)^3 = (0 \| 3 \| 1) = 8.$

Check:

$(C^{\rightarrow}_{MNS})^3 = 32^3 = 32768 = 8(\text{mod}\,70) = C^{\rightarrow}_{MNS} \times C^{\rightarrow}_{MNS} \times C^{\rightarrow}_{MNS} = (0 \| 2 \| 4) \times$
$\times (0 \| 2 \| 4) \times (0 \| 2 \| 4) = (0 \| 3 \| 1) = 8.$

$$C^{3}_{MNS} = (C^{3}_{MNS})^{\rightarrow} - \frac{1}{2} D, \ \ (-3)^3 = 8 - 35 = -27, \ \ (-3)^3 = -27.$$

Based on the developed method, let's consider an algorithm for squaring the residues of integers modulo p_k MNS. This algorithm is often used in cryptographic calculations [33-35]. Let it be necessary to determine the value $c_k^2(\text{mod}\, p_k)$, where: c_k, p_k are natural numbers and the condition $0 \le c_k \le p_k - 1$ is satisfied. First, let's show that the following mathematical ratio is fulfilled:

$$c_k^2(\text{mod}\, p_k) \equiv (p_k - c_k)^2 \text{ mod } p_k. \qquad (16)$$

Indeed, the number c_k^2 can be represented in the form $c_k^2 = i \cdot p_k + \alpha$ given that $0 \le \alpha \le p_k - 1;\ i = 0, 1, 2, ...,$ i.e. $c_k^2 \equiv \alpha \text{ mod } p_k$. Then $(p_k - a_i)^2 = p_k^2 - 2 \cdot p_k \cdot c_k + c_k^2 = \ = p_k^2 - 2 \cdot p_k \cdot c_{ki} + i \cdot p_k + \alpha.$ In this case $(p_k^2 - 2 \cdot p_k \cdot c_{ki} + i \cdot p_k + \alpha) \equiv \alpha \text{ mod } p_k$. Equality (16) is valid for both even and odd values of the modulus p_k of the MNS.

Analytical ratio (16) is an algorithm for squaring the residues of integers modulo MNS. Let's consider three possible options for the practical implementation of this algorithm.

First option. The modulus value $p_k = 2 \cdot n + 1$ ($n = 0, 1, 2, ...$) is an odd number. For this case, the process diagram for implementing the operation $C^2(\text{mod}\, p_k)$ is directly based on mathematical ratio (16).

Second option. The modulus value $p_k = 2 \cdot n$ and the value $\frac{p_k}{2}$ are even numbers. In this case, the value $\frac{p_k}{2}$ is an integer and, therefore,

$\left(\frac{p_k}{2}\right)^2 = \frac{p_k}{4} \cdot p_k \equiv 0(\text{mod } p_k)$. In this case, the algorithm for implementing the operation of squaring the residues of integers modulo MNS is determined by the following ratio:

$$\left(\frac{p_k}{2}\right)^2 \equiv 0(\text{mod } p_k). \tag{17}$$

Third option. The modulus value $p_k = 2 \cdot n$ is an even number, and the value $\frac{p_k}{2}$ is an odd number. The proposed algorithm for squaring the residues of integers modulo MNS is based on the use of the following ratio:

$$\left(\frac{p_k}{2}\right)^2 \equiv \frac{p_k}{2}(\text{mod } m_i). \tag{18}$$

Indeed, mathematical ratio (18) can be easily represented in the form:

$$\frac{p_k}{2} \cdot \left(\frac{p_k}{2} - 1\right) \equiv 0(\text{mod } \frac{p_k}{2} \cdot 2). \tag{19}$$

From number theory it is known that the comparability of $C \equiv D(\text{mod } p_k)$ of two numbers C and D modulo p_k is equivalent to the divisibility of the number $C - D$ by modulo p_k [36, 37].

From ratio (19) it follows that the number $\frac{p_k}{2} \cdot \left(\frac{p_k}{2} - 1\right)$ is divided by the module $p_k = \frac{p_k}{2} \cdot 2$. Indeed, the first factor (multiplier) $\frac{p_k}{2}$ of the product (19) is divided by $\frac{p_k}{2}$, and the second factor $\frac{p_k}{2} - 1$ is divided by two, since by condition $\frac{p_k}{2}$ is an odd number. Thus, the validity of comparison (18) is shown.

Thus, the main result of this research is that, based on the mathematical model (8) of the process of squaring numbers (integers) modulo, a mathematical model (9) of the process of exponentiating the residue of numbers modulo MNS to the power of a natural number, both positive and negative numeric domains. Using this SMR, a method for exponentiating numbers in the MNS in all numeric domain and an algorithm for squaring the residues of integers modulo MNS has been developed. The method and algorithm are based on the use of properties of the MNS $(-C) \bmod p_k \equiv (p_k \cdot i - C) \bmod p_k, \quad (i = 1, 2, 3, ...)$,

$c_k^2 (\bmod p_k) \equiv (p_k - c_k)^2 \bmod p_k$, $\left(\frac{p_k}{2}\right)^2 \equiv 0(\bmod\ p_k)$ or

$\left(\frac{p_k}{2}\right)^2 \equiv \frac{p_k}{2} (\bmod m_i),$ by using the tabular principle of data processing [38-40]. The results obtained are important and can be used for technical implementation of the operation of exponentiation in computer systems for processing integer data, operating both in the MNS and in conventional binary positional number systems.

6. Conclusion

The procedure for realization the operation of exponentiating integers in the MNS in a positive numeric domain has been researched. As the operation of exponentiating numbers is one of the key operations of many algorithms and protocols that are used in modern CS (programming, cryptography, optimization algorithms, etc.) and there is no effective implementation of this operation, especially for negative numbers. Two options for representing numbers in the MNS are considered, in all numeric domain. The first option is as follows: the original number in the MNS has an additional two sign bits. These bits symbolize the sign of the number in the MNS. The second option is as follows: for perform the process of realizing the operation of exponentiating numbers by modulo MNS in all numeric domain, it is supposed to represent the original number in a modular structure. The research developed a method for exponentiating numbers in the MNS, both in positive and negative numeric domain. This method is based on the use of a synthesized mathematical model in an analytical ratio, which is a generalized system of mathematical ratios of the process of exponentiating numbers modulo MNS. The development of the method was carried out by applying a special coding of numbers in the MS using the tabular principle of data processing. The use of the tabular principle increases the speed of implementation of the operation of exponentiating integers, which improves the performance of the CS in the MNS. The result of the developed method is presented in the form of examples of the operation of exponentiating numbers represented in the MNS. An analysis of the solution of examples showed the practical value of the developed method. Based on the obtained method, algorithm for implementing the operation of squaring the residues of integers by an arbitrary modulo MNS was received, in accordance with which devices for their implementation were synthesized. These technical devices, for which Ukrainian patents have been received, are recommended for use in the practical implementation of CS components operating in both the MNS and the PNS.

7. References

[1] S. Hamedi-Hagh, Error-Free Calculation of Total Referred Noises in Electronic Circuits, in: 2019 16th International Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD), Lausanne, Switzerland, 2019, pp. 97–100. DOI: 10.1109/SMACD.2019.8795229.

[2] M. Technau, Remark on the Farey fraction spin chain, (2023). DOI: 10.48550/arXiv.2304.08143.

[3] I. Ya. Akushskii, D. I. Yuditskii, Machine Arithmetic in Residual Classes, Sov. Radio, Moscow, 1968.

[4] Y.M. Nikolaychuk (Ed.), Specialized computer technologies in informatics, TernogrMS, Ternopil, 2017.

[5] A. I. Kornilov, M. Yu. Semenov, V. S. Kalashnikov, Methods for Hardware Optimization of Adders for Two Operands in the System of Residual Classes, 1st. ed., Electronics, Moscow, 2017.

[6] A. Sasaki, The Basis for Implementation of Adidive Operations in the Residue Number System, in: IEEE Transactions on Computers C-17, (1968) 1066–1073. DOI:10.1109/TC.1968.226466.

[7] V. Krasnobayev, A. Kuznetsov, A. Yanko, S. Koshman, A. Zamula and T. Kuznetsova, Data processing in the system of residual classes, ASC Academic Publishing, Minden, Nevada, USA, (2019). ISBN: 978-0-9989826-6-3 (Hardback), ISBN: 978-0-9989826-7-0 (Ebook).

[8] Exponentiation by squaring, 2001. [Online]: https://en.wikipedia.org/wiki/Exponentiation_by_squaring

[9] A. Cui, H. Zhao, X. Zhang, B. Zhao, Z. Li, Power system real time data encryption system based on DES algorithm, in: 2021 13th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), IEEE Press, Beihai, 2021, pp. 220–228. DOI: 10.1109/ICMTMA52658.2021.00056.

[10] V. Mainanwal, M. Gupta, S. K .Upadhayay, Zero Knowledge Protocol with RSA Cryptography Algorithm for Authentication in Web Browser Login System (Z-RSA), in: 2015 Fifth International Conference on Communication Systems and Network Technologies, IEEE Press, Gwalior, 2015, pp. 776–780. DOI: 10.1109/CSNT.2015.90.

[11]Ch. Pakkiraiah, R. Satyanarayana, Design and FPGA Realization of an Energy Efficient Artificial Neural Modular Exponentiation Architecture. Computing, Communication and Learning, (2023) 115–126. DOI: 10.1007/978-3-031-21750-0_10.
[12]D. Schinianakis, T. Stouraitis, Number Systems in Cryptography: Design, Challenges, Robustness, 2016. DOI: 10.1007/978-3-319-14971-4_4.
[13]I. Yakymenko, M. Kasianchuk, O. Gomotiuk, G. Tereshchuk, S. Ivasiev, P. Basistyi, Elgamal cryptoalgorithm on the basis of the vector-module method of modular exponentiation and multiplication, in: 2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), Lviv-Slavske, 2020, pp. 926–929. DOI: 10.1109/TCSET49122.2020.235572.
[14]V. R. Kolagatla, M. J, S. Darbar, D. Selvakumar, S. Saha, A Randomized Montgomery Powering Ladder Exponentiation for Side-Channel Attack Resilient RSA and Leakage Assessment, in: 2021 25th International Symposium on VLSI Design and Test (VDAT), Surat, 2021, pp. 1–5. DOI: 10.1109/VDAT53777.2021.9601132.
[15]K. El Makkaoui, Y. Lamriji, I. Ouahbi, O. Nabil, A. Bouzahra, A. Beni-Hssane, Fast Modular Exponentiation Methods for Public-Key Cryptography, in: 2022 5th International Conference on Advanced Communication Technologies and Networking (CommNet), Marrakech, Morocco, 2022, pp. 1–6. DOI: 10.1109/CommNet56067.2022.9993851.
[16]K. Venkatesh, K. Pratibha, S. Annadurai, L. Kuppusamy, Reconfigurable Architecture to Speed-up Modular Exponentiation, in: 2019 International Carnahan Conference on Security Technology (ICCST), Chennai, 2019, pp. 1–6. DOI: 10.1109/CCST.2019.8888401.
[17]D. An, Y. Li, S. Zhang, J. Lu, Securely Outsource Modular Exponentiations With Single Untrusted Cloud Server, IEEE Access, 8 (2020), 200764–200773. DOI: 10.1109/ACCESS.2020.3033306.
[18]Y. Nanjo, M. Shirase, Y. Kodera, T. Kusaka, Y. Nogami, A Construction Method of Final Exponentiation for a Specific Cyclotomic Family of Pairing-Friendly Elliptic Curves with Prime Embedding Degrees, in: 2021 Ninth International Symposium on Computing and Networking (CANDAR), Matsue, Japan, 2021, pp. 148–154. DOI: 10.1109/CANDAR53791.2021.00028.
[19]W. Tan, S.-W. Chiu, A. Wang, Y. Lao, K.K. Parhi, PaReNTT: Low-Latency Parallel Residue Number System and NTT-Based Long Polynomial Modular Multiplication for Homomorphic Encryption, 2023. DOI: 10.48550/arXiv.2303.02237.
[20]V. Krasnobayev, S. Koshman, A. Yanko, A device for raising numbers represented in the residual class to the power of a natural number, 2014. Patent for useful model No. 95060, Filed June 18th., 2014, Issued December 10th., 2014.
[21]V. Krasnobayev, A. Zamula, S. Rassomakhin, A. Yanko, A device for squaring numbers in the residual class system, 2014. Patent for useful model No. 129249, Filed April 23th., 2014, Issued October 25th., 2014.
[22]V. Krasnobaev, A. Kuznetsov, V. Babenko, M. Denysenko, M. Zub, V. Hryhorenko, The Method of Raising Numbers, Represented in the System of Residual Classes to an Arbitrary Power of a Natural Number, in: 2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON), IEEE Press, Lviv, 2019, pp. 1133–1138. DOI: 10.1109/UKRCON.2019.8879793.
[23]V. Krasnobayev, A. Kuznetsov, I. Lokotkova, A. Kiian, T. Kuznetsova, Techniques for Raising the Remainder to a Power in the System of Residual Classes, in: 2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT), IEEE Press, Kyiv, 2020, pp. 145–150. DOI: 10.1109/DESSERT50317.2020.9125049.
[24]Z. Xiaoqiang, Z. Xueheng, Image encryption algorithm based on the Matryoshka transform and modular-inverse matrix, 2023. DOI: 10.21203/rs.3.rs-2663096/v1.

[25] S. Banerjee, S. Chakraborty, N. Dey, A. Kumar Pal, R. Ray, High Payload Watermarking using Residue Number System. International Journal of Image, Graphics and Signal Processing 7 (2015) 1–8. DOI: 10.5815/ijigsp.2015.03.01.
[26] K.A. Gbolagade, S.D. Cotofana, Residue Number System operands to decimal conversion for 3-moduli sets, in: 2008 51st Midwest Symposium on Circuits and Systems, IEEE, 2008, pp. 222–224. DOI: 10.1109/MWSCAS.2008.4616918.
[27] B. Parhami, On equivalences and fair comparisons among residue number systems with special moduli, in: 2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers, IEEE, 2010, pp. 1690–1694. DOI: 10.1109/ACSSC.2010.5757827.
[28] K. Phalakarn, A. Surarerks, Alternative Redundant Residue Number System Construction with Redundant Residue Representations, in: 2018 3rd International Conference on Computer and Communication Systems (ICCCS), IEEE, 2018, pp. 457–461. DOI: 10.1109/CCOMS.2018.8463305.
[29] D. Younes, P. Steffan, Efficient image processing application using residue number system, in: Proceedings of the 20th International Conference Mixed Design of Integrated Circuits and Systems - MIXDES 2013, IEEE, Gdynia, 2013, pp. 468–472.
[30] V. Yatskiv, T. Tsavolyk, N. Yatskiv, The correcting codes formation method based on the residue number system, in: 2017 14th International Conference The Experience of Designing and Application of CAD Systems in Microelectronics (CADSM), IEEE, 2017, pp. 237–240. DOI: 10.1109/CADSM.2017.7916124.
[31] D.I. Popov, A.V. Gapochkin, Development of Algorithm for Control and Correction of Errors of Digital Signals, Represented in System of Residual Classes, in: 2018 International Russian Automation Conference (RusAutoCon), IEEE, Sochi, 2018, pp. 1–3. DOI: 10.1109/RUSAUTOCON.2018.8501826.
[32] H. Bombín, C. Dawson, Ye-Hua Liu, N. Nickerson, F. Pastawski, S. Roberts, Modular decoding: parallelizable real-time decoding for quantum computers, 2023.
[33] Yu.I. Gorbenko, M.V. Yesina, V.A. Ponomar, I.D. Gorbenko, E.Yu. Kapt'ol, Scientific and methodological bases of analysis, evaluation and results of comparison of existing and promising (post-quantum) asymmetric cryptographic primitives of electronic signature, protocols of asymmetric encryption and key encapsulation protocols, Radiotekhnika, 212 (2023) 42–65. DOI: 10.30837/rt.2023.1.212.05.
[34] O. Tsentseria, K. Hleha, A. Matiyko, I. Samoilov, The state of standardization of post-quantum crypto-algorithms at the global level, Automation of technological and business processes, 15(2), (2023) 66–71. DOI: 10.15673/atbp.v15i2.2527.
[35] X. Ma, M. Wang, Z. Kang, Distributed Implementation of SM4 Block Cipher Algorithm Based on SPDZ Secure Multi-party Computation Protocol, Data Science. ICPCSEE 2023. Communications in Computer and Information Science, 1879, (2023) 279–297. DOI: 10.1007/978-981-99-5968-6_20.
[36] V. Platonov, A. Rapinchuk, I. Rapinchuk, Algebraic Number Theory, Algebraic Groups and Number Theory, (2023) 1-52. DOI:10.1017/9781139017756.004.
[37] T. Ward, Ergodic Theory: Interactions with Combinatorics and Number Theory, Ergodic Theory, (2023) 577–596. DOI: 10.1007/978-1-0716-2388-6_181.
[38] V. A. Krasnobayev, A. S. Yanko, D. M. Kovalchuk, Methods for tabular implementation of arithmetic operations of the residues of two numbers represented in the system of residual classes, Radio Electronics, Computer Science, Control, 4, (2022) 18–27. DOI: 10.15588/1607-3274-2022-4-2.

[39]V. A. Krasnobayev, A. S. Yanko, S.A. Koshman, Method of tabular realization of arithmetic operations in the system of residual classes, Computer science and cybersecurity, 3 3 (2016) 28–35.
[40]D. Panfilo et al., A Deep Learning-Based Pipeline for the Generation of Synthetic Tabular Data, IEEE Access, 11, (2023) 63306–63323. DOI: 10.1109/ACCESS.2023.3288336.

UDC 511.334:511-33

МЕТОД ОБЧИСЛЕННЯ ПІДНЕСЕННЯ ДО СТЕПЕНЯ ЗА МОДУЛЕМ ДОДАТНИХ І ВІД’ЄМНИХ ЦІЛИХ ЧИСЕЛ

Dr.Sci. Краснобаєв В.[1], Ph.D. Янко А.[2], Ph.D. Мартиненко А.[3], Ковальчук Д.[4]
[1,4]Харківський національний університет імені В. Н. Каразіна, Україна
[2,3]Національний університет «Полтавська політехніка імені Юрія Кондратюка», Україна
E-mail: [1]v.a.krasnobaev@gmail.com, [2]al9_yanko@ukr.net, [3]martynenko@pntu.edu.ua [4]kovalchuk.d.n@ukr.net

Анотація. *Численні публікації останніх років свідчать про те, що перспективи створення високошвидкісних комп'ютерних систем (КС) на основі використання модульної системи числення (МСЧ) відкривають широкі можливості для використання КС з високим ступенем розпаралелення обробки цілочисельних даних. Застосування основних властивостей МСЧ і можливості використання табличного принципу обробки даних значно підвищує швидкість виконання цілочисельних арифметичних операцій у порівнянні з традиційною двійковою позиційною системою числення. Дослідження в цій галузі показують ефективність використання МСЧ для збільшення швидкості виконання, окрім основних арифметичних операцій (додавання, віднімання та множення) над цілими числами, операції піднесення цілих чисел до степеня. Однак досі не існує ефективних методів піднесення чисел до степеня в МСЧ у всій числовій області (додатній і від’ємній). Тому в цьому дослідженні розроблено систему математичних співвідношень, що описують досліджуваний процес, на основі якої розроблено метод піднесення чисел до степеня, який, на відміну від відомих, може бути реалізований у від’ємній числовій області. На основі отриманого методу отримано алгоритм піднесення чисел до степеня в МСЧ, за якими синтезовано пристрої для їх реалізації.*

Ключові слова: *арифметична операція, комп'ютерна система, система математичних співвідношень, метод піднесення чисел до степеня, модульна система числення, модульна структура, позиційна система числення, табличний код множення.*

Section 7. Intelligent systems and data analysis

UDC 004:37:001:62

DIAGNOSIS INTELLECTUALIZATION OF COMPLEX TECHNICAL SYSTEMS

Dr.Sci. Vychuzhanin V., Vychuzhanin A.

National University "Odessa Polytechnic"
ORCID: 0000-0002-6302-1832, 0000-0001-8779-2503
E-mail: v.v.vychuzhanin@op.edu.ua

Abstract. *The article presents the results of developing a model for diagnosing a ship complex technical system with incomplete data and its implementation in an intelligent system for assessing the risk of failures of subsystems, components, intercomponent links, which allows obtaining a priori information about the technical condition of a complex system. Types of technical condition of subsystems, components, intercomponent connections are determined on the basis of diagnostic features of a complex system using the example of a ship power plant to assess the risk of their failures. Predicting the type of technical state of a complex technical system was carried out using a posteriori inference in Bayesian belief networks. The studies presented in the article assessed the risk of failures as a result of the use of an intelligent system for diagnosing and predicting the risk of failures of a ship complex technical system. The model for diagnosing and predicting the risk of failures of subsystems, components, interconnections can be considered as a conceptual model of an intelligent system for diagnosing and predicting the risk of failures of complex technical systems on network infrastructures, which has a relative insensitivity to incomplete technological data.*
Keywords*: technical condition; complex technical system; risk of failure; diagnostics; forecasting; intelligent system; Bayesian belief network; insensitivity to incomplete data.*

1. Introduction

Complex technical systems (CTS), having structural and functional diversity, differ in the principles of operation and consist of numerous interconnected and interdependent subsystems, components with complex intersystem, intercomponent links [1, 2]. The increase in the complexity of the composition of ship CTS affects the growth of system failures, which in turn is accompanied by an increase in repair work or CTS components replacement. Intellectualization of diagnostics and forecasting of the technical condition (TC) of ship complex systems makes it possible to extend the operation time of such CTS. This article is devoted to the development and application of artificial intelligence methods for evaluating the TC of subsystems (TS), components (CM), intersystem (IS) and CTS intercomponent communications (ICC) in order to prevent their possible failures.

2. Description of Problem

Diagnostics and prediction of TC helps to reduce the CTS risk of failures of subsystems, components, intersystem and intercomponent connections at their operation stage [3, 4, 5, 6]. Diagnosis (evaluation) and prediction of the technical state of the CTS should take into account the specifics of systems that are often operated under uncertain conditions of the external and internal environment, with unspecified CTS regulatory parameters values and have relative insensitivity to incomplete system components technological data [7]. The CTS reliability can be assessed by the results of diagnosing the TC, and the prediction of changes in the TC makes it possible to operate the systems until signs of a dangerous decrease in reliability appear, while excluding premature dismantling of components and assemblies, as well as performing other labor-intensive work that is often of dubious usefulness for the CTS reliable operation.

To successfully solve the problem of ensuring the ship CTS reliability it is necessary to remove a number of uncertainties. Such uncertainties include incomplete data on external, internal impacts on systems and on the state of such systems. The removal of uncertainties can be based on solving problems of assessing the risk of CTS failures and its prediction with relative insensitivity to incomplete technological data FS, FC, FIC and FI [8, 9, 10]. Traditional automation of ship equipment includes, in addition to monitoring parameters and controlling installations and mechanisms, also equipping them with complexes that allow us to create hierarchical distributed integrated systems for diagnosing and predicting complex systems TC [1, 11, 12].

Diagnostic and forecasting systems should: constantly carry out diagnostics of the ship CTS TC of functional FS, FC, FIC and FI; analyze trends in changes in the technical equipment of systems; perform failover and provide TS prediction. To implement such a technology, appropriate algorithmic and software tools are required. The diagnostic algorithms used, as a rule, are based on the tolerance control of individual diagnostic parameters.

However, the analysis and integral assessment of FS, FC, FIC and FI TC, the development of control actions in most cases is carried out by ship operators based on heuristic rules. At the same time, the volume of measuring and diagnostic information, the number of connections, dependencies of systems diagnostic TC features and types can be significant.

In theory, engineering practice, various methods are used to assess the risk of failures FS, FC, FIC and FI CTC. An example of the application of risk theory is the logical development of probabilistic analysis to assess CTS failures risk [13]. Advantages of probabilistic analysis: the full range of accident scenarios and consequences of FS, FC, FIC and FI failures is analyzed, and not only design basis accidents; balanced approach; an objective assessment of the accident rate is used; accounting for the interdependence between CTS subsystems and components in an explicit form. With a probabilistic approach, the reliability level is selected depending on the possible consequences in case of damage (failure) of FS, FC, FIC and FI systems. In this regard, the assessment of the risk of ship CTS failures lies in their damage unacceptable probability.

However, the negative consequences of a failure in systems are often taken into account intuitively, implicitly, by taking certain failure-free operation probability values or system components safety factor.

Intellectualization of automated diagnostic systems involves solving a number of interrelated tasks of a structural, functional, informational and organizational nature, which should be provided at the stage of designing diagnostic systems for TC CTS [6]. Expert systems are widely used to automatically analyze data and issue recommendations to prevent possible failures. These systems can be integrated with control and monitoring a system, which allows us to quickly respond to changes in the parameters of the CTS operation and take measures to ensure its safety and reliability.

In artificial intelligence, knowledge representation models are actively developing - Bayesian Belief Networks (BBN), used to diagnose the TC of complex systems [14, 15, 16].

One of the BBNs advantages for vehicle diagnostics is their ability to work with uncertain and incomplete CTS process data. BBN can be applied to assess the risk of failures in CTS, providing data and knowledge integration to assess the likelihood of various failure scenarios and their consequences. By identifying critical components, evaluating maintenance strategies, and supporting regulatory compliance, BBNs can help ensure safe and reliable CTS operations. Thus, the problems associated with ensuring the reliable operation of ship CTS require further improvement and the search for new methods, models and algorithms aimed at promptly detecting emergency conditions of equipment, at solving the problems of diagnosing and predicting system failures risk under conditions of relative insensitivity to incomplete data on FS, FC, FIC and FI.

Since all modern ships must be equipped with automation systems for technical means using artificial intelligence technologies, the introduction of approaches based on such methods, models and algorithms should help ensure ship’s CTS reliable operation.

That is, taking into account the existing problems in ensuring reliability during the operation of CTS, the intellectualization of diagnostics and predicting ship’s CTS failures risk by diagnostic features is a significant direction that allows us to influence the safety and reliability of systems and is an urgent task.

Statement of the problem: to substantiate the forecast failures risk of FS, FC, FIC and FI by intellectualizing the assessment ship complex systems TC by diagnostic features. The purpose of the work: ensuring the reliability and safety of ship CTS by reducing the risk of failures FS, FC, FIC and FI.

3. Model diagnostics technical condition intellectualization and prediction ship complex systems failures risk

In a formalized form, BBN for diagnosing the TC and predicting the risk of ship CTS failures contains an acyclic directed graph G, a set of vertex variables and directed links between them. A formalized generalized model for the intellectualization of TC diagnostics and predicting failures risk of FS, FC, FIC and FI by diagnostic features can be described as follows:

$$< G, S(C), I_S(I_C), R_{S(C)}, R_{I_S(I_C)}, L >, \quad (1)$$

where: $S(C)$, set FS (FC);

$I_S(I_C)$, set FIC (FI);

$R_{S(C)}, R_{I_S(I_C)}$, set of failures risk diagnostic assessments FS (FC) и FIC (FI) CTS;

L - mapping relationships between sets $S(C), I_S(I_C)$ and $R_{S(C)}, R_{I_S(I_C)}$, based on the fault tree of the CTS diagnostic mode.

TC FS (FC) and FIC (FI) sets with $S(C), I_S(I_C)$ is determined based on the failure tree (Failure Tree), presented as a set failures risk of FS, FC, FIC and FI. The tree consists of failure risk sequences, which are a multi-level graphological structure diagnostic model causal relationships, obtained by tracking failures in the reverse order of the structure, in order to find occurrence their possible causes. The advantage of a fault tree over other failure scoring models is that failure analysis is limited to identifying only those FS, FC, FIC, and FI of the system and the events that lead to a particular system failure or crash. The fault tree allows us to identify all the paths leading to the failure of the CTS, and provides the determination of the minimum number of combinations of events that cause the system to fail [17]. The set FS (FC) of the CTS, taking into account their hierarchical levels, is determined

$$S(C) = \{\upsilon_{n_{s(c)}}^{<m_{s(c)}>} \mid n_{s(c)} = \overline{1, N_{S(C)}}; m_{s(c)} = \overline{1, M_{S(C)}}\}, \qquad (2)$$

where $\upsilon_{n_{s(c)}}^{<m_{s(c)}>}$, each FS (FC) condition;

$n_{s(c)}$, FS (FC) number;

$m_{s(c)}$, hierarchical level number FS (FC);

$N_{S(C)}$, FS (FC) value;

$M_{S(C)}$, hierarchical level number value FS (FC).

The state of each FS and FC CTC is expressed as:

$$\upsilon_{n_{S(C)}}^{<m_{S(C)}>} = \{W^0_{\upsilon_{S(C)n(m)}}, W^f_{\upsilon_{S(C)n(m)}}, a_{\upsilon_{in_{S(C)n(m)}}}, a_{\upsilon_{on_{S(C)n(m)}}}\}, \qquad (3)$$

where $W^0_{\upsilon_{S(C)n(m)}}$, full working capacity FS (FC);

$W^{f}_{\upsilon_{S(C)n(m)}}$, performance FS (FC) at different degrees of their losses;

$a_{\upsilon_{in_{S(C)n(m)}}}, a_{\upsilon_{on_{S(C)n(m)}}}$, incoming and outgoing FS (FC);

in, on, sequence number of incoming and outgoing FIC, FI.

Operability of FS and FC at different degrees of their losses:

$$W^{f}_{\upsilon_{S(C)n(m)}} = \{W_f^{<n_{s(c)}, m_{s(c)}>} \mid f = \overline{0,1}; n_{s(c)} = \overline{1, N_{S(C)}}; m_{s(c)} = \overline{1, M_{S(C)}}\}, \quad (4)$$

In (4) f =0 is the correct state of the STS, f=1 is the failure of the CTS. The set of FIC and FI CTS is determined by:

$$I_{S(C)} = \{\omega^{<a,b,z,q>}_{I_{S(C)}} \mid a = \overline{1, A}; b = \overline{1, B}; z = \overline{1, Z}; q = \overline{1, Q}\}, \quad (5)$$

where $\omega^{<a,b,z,q>}_{I_{S(C)}}$, state of each FIC (FI);

a, FIC number;
z, FI number;
b, hierarchical level number FIC;
q, hierarchical level number FI;
A, FIC value;
Z, FI value;
B, hierarchical level value FIC;
Q, hierarchical level value FI.

State of each FIC and FI:

$$\omega^{<a,b,z,q>}_{I_{S(C)}} = \{W^{0}_{\omega_{I_{S(C)a(z)}}}; W^{f}_{\omega_{I_{S(C)a(z)}}}\}, \quad (6)$$

where $W^{0}_{\omega_{I_{S(C)a(z)}}}$, full working capacity FIC (FI);

$W^{f}_{\omega_{I_{S(C)a(z)}}}$, operability FIC (FI) at different degrees of their loss.

Operability of FIC and FI at different degrees of their loss:

$$W^{f}_{\omega_{I_{S(C)_{a(z)}}}} = \{W_f^{<a,z>} \mid f = \overline{0,1}; a = \overline{1,A}; z = \overline{1,Z};\} \quad (7)$$

$n(m)$ FS (FC) failure risk:

$$R_{S(C)_{n(m)}} = D_{S(C)_{n(m)}} \cdot P_{S(C)_{n(m)}}(t) \quad (8)$$

$a(z)$ FIC (FI) failure risk:

$$R_{I_{S(C)_{a(z)}}} = D_{I_{S(C)_{a(z)}}} \cdot P_{I_{S(C)_{a(z)}}}(t) \quad (9)$$

Modeling assumptions and limitations are that FS and FC CTSs can have a failure risk level distributed based on the Harrington desirability function [18]: 0 - 0.2 - minimal (the consequences of failure are minimal); 0.2 - 0.37 acceptable (consequences of failure are insignificant); 0.37 - 0.63 - maximum (consequences of failure are significant); 0.63 - 1 - critical. An available source for the reliability statistics FS, FC, FIC and FI of ship CTS when choosing the values of the conditional probabilities of their failures to determine failures risk is the OREDA offshore database [19]. In the database, conditional probabilities correspond to the exponential distribution law for the time between failures FS, FC, FIC and FI, whose resource is installed before the end of the normal operation period. CTS with a set of FS, FC, FIC and FI can be classified as "non-aging" systems, since they operate only in the area with the failure rate λ(t)=λ=const. The failure rate is:

$$\lambda(t) = \frac{\alpha \cdot \exp(-\alpha T_0)}{\exp(-\alpha T_0)} = \alpha \quad (10)$$

where α – distribution parameter, taken according to the test results equal to $\alpha \approx 1/\widehat{T}_o$, $\widehat{T}_o$ – mean time to failure estimate.

Mean time to failure:

$$T_0 = \int_0^{\infty} P(t)dt = 1/\lambda \quad (11)$$

The failure probability FS and FC CTS, taking into account the number of failures of a certain subsystem (component), can also be determined:

$$P_{S(C)_{n(m)}} = \frac{v_{S(C)_{n(m)}}}{\tau} \quad (12)$$

where $P_{S(C)_{n(m)}}$, failure probability $n(m)$ FS (FC);

$v_{S(C)_{n(m)}}$, number of failures $n(m)$ FS (FC);

$\tau = 10^6$ hours - the period of statistical testing.

The probability of failure FIC and FI CTS, taking into account the number of failures of a certain intersystem (intercomponent) connection, can be determined:

$$P_{I_{S(C)_{a(z)}}} = \frac{v_{I_{S(C)_{a(z)}}}}{\tau} \tag{13}$$

where $P_{I_{S(C)_{a(z)}}}$, conditional failure probability $a(z)$ FIC and FI;

$v_{I_{S(C)_{a(z)}}}$, number of failures $a(z)$ FIC and FI.

Quantitative assessment of damage FS, FC from failure of a subsystem (component) to determine the risk of failure:

$$D_{S(C)_{n(m)}} = \frac{Nu_{n(m)} \cdot (C_{n(m)} + Ci_{n(m)} + Cd_{n(m)})}{Nf_{n(m)}}; \tag{14}$$

where $Nu_{n(m)}$, number of unrecoverable failures FS (FC);

$C_{n(m)}$, FS (FC) price;

$Ci_{n(m)}$, FS (FC) installation cost;

$Cd_{n(m)}$, FS (FC) disposal cost;

$Nf_{n(m)}$, number of failures FS (FC);

$d_{s(c)_{n(m)}}$, damage from abandonment FS (FC).

Quantification of FIC and FI damage from failure $a(z)$ FIC and FI:

$$D_{I_{S(C)_{a(z)}}} = \frac{Nu_{a(z)} \cdot (C_{a(z)} + Ci_{a(z)} + Cd_{a(z)})}{Nf_{a(z)}}; \tag{15}$$

where $Nu_{a(z)}$, number of unrecoverable failures FIC (FI);

$C_{a(z)}$, price FIC (FI);

$Ci_{n(m)}$, installation cost FIC (FI);

$Cd_{a(z)}$, disposal cost FIC (FI);

$Nf_{a(z)}$, number of failures FIC (FI);

$d_{i_{s(c)_{a(z)}}}$, damage from abandonment FIC (FI).

The verbal form is used to describe the category of damage from refusal FS (FC). To compare numerical estimates for different classes of damage, the Harrington scale is used [18]: “Insignificant damage” - 0.1· Dcrit; "Damage is insignificant" - 0.29· Dcrit; "Damage of medium significance" - 0.51· Dcrit; “Significant damage” – 0.72· Dcrit; "Damage Critical" –1· Dcrit. According to the established conditional failure probabilities and damages from failures FS, FC, FIC and FI according to (8), (9), their risk of failures is determined. The initial data for constructing an intellectualization model for assessing the technical condition and predicting the risk of failures of complex systems on the example of a ship power plant (SPP) based on dynamic are: PPS principle operation scheme; failure probabilities FS, FC, FIC and FI CTS. The set of TS FS, FC, FIC and FI CTS is determined based on the failure tree, presented as a set of their failure risk (Fig. 1). Symbols of subsystems, components of the SPP in BBN: Input element - IE; Fire fighting system - FFS; Compressed air system - CAS; Manual control of the main engine - MCME; Control system - CS; Remote automated control system of the main engine - RACSME; Intermediate component - P1; Ship power plant - SPP; Main engine - ME; Ballast drainage system - BDS; Emergency drive propulsion and steering complex - ED PSC; Control system for propulsion and steering complex -CSPSC; Boiler plant - BP; Transfer of power from the main engine to the propeller - TPMEP; Intermediate component - P2; Propulsion and steering complex - PSC; Output component - EXIT. Tab. 1 reflects the correspondence of symbols on the fault tree S and FS, FC BBN. The structure of the BBN SPP, shown in Fig. 2, is a multi-level subsystem location system, consisting of 13 subsystems, 7 levels with the addition of specialized intermediate nodes P1 and P2, providing the implementation of a multi-level network structure. For the subsystems of the upper level SPP BBN structure, conditional failure probabilities are specified, taking into account the influence of subsystems of a lower hierarchical level on subsystems of a higher level.

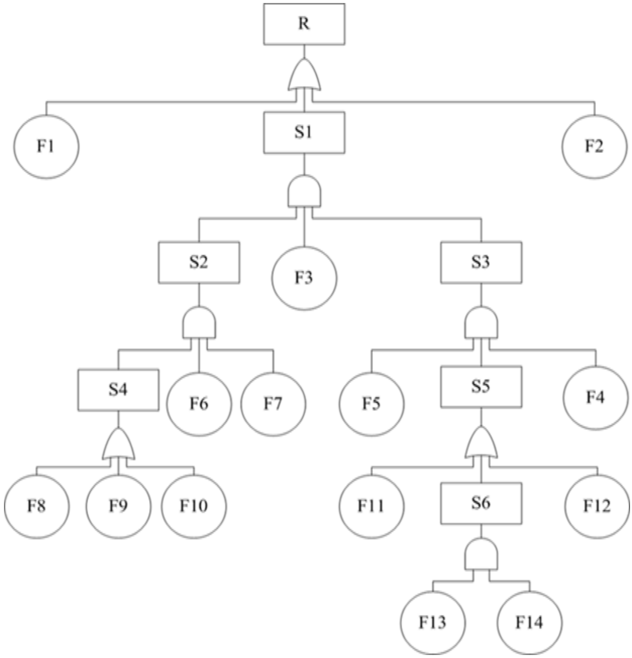


Fig. 1. Fault tree of subsystems (components), intersystem (intercomponent) connections of the SPP

Table 1.

Correspondence table S and subsystems (components) BBN

Designation	Event characteristics
S1	Violation of the IE element
S2	Violation of the FFS, CAS, MCME **elements**
S3	Violation of the RACSME, P1, SPP elements
S4	Violation of the CS, BDS, BP elements
S5	Violation of the ME, ED_PSC, CSPSC elements
S6	Violation of the TPMEP, P2, PSC elements

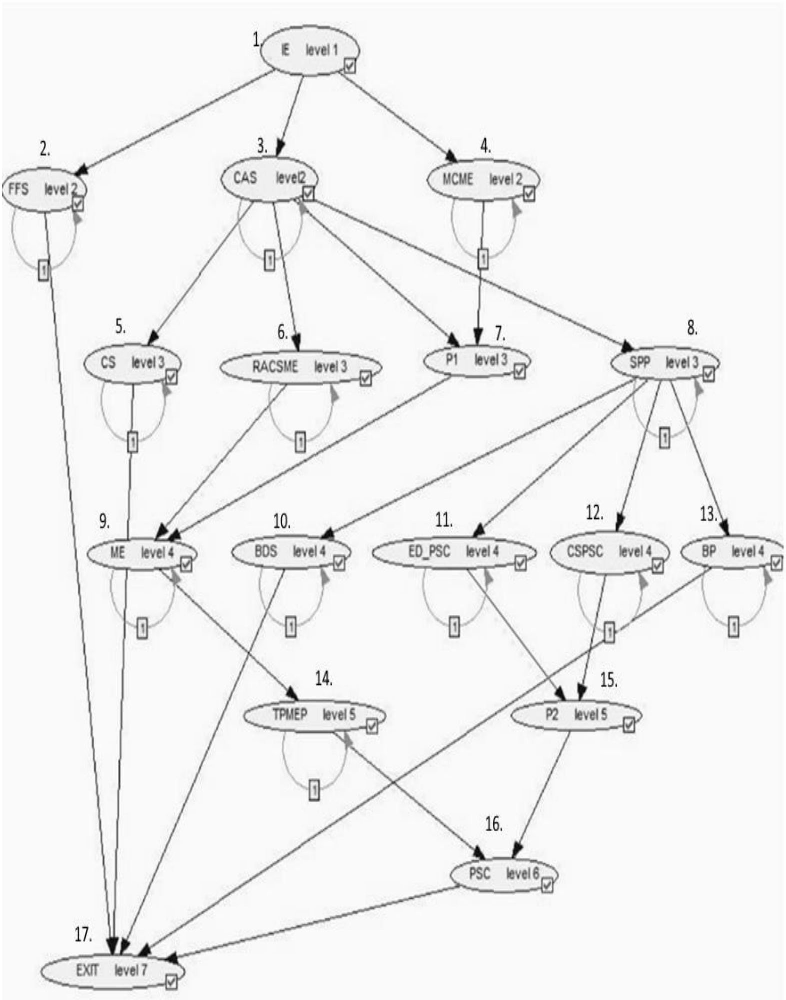


Fig. 2. Structure of BBN SPP

As an BBN example for interconnected power plant units (Fig. 2) IE, CAS, SPP and interconnections IE-CAS, CAS - SPP, sets of failure risk at the initial time and taking into account the dynamics of technical conditions over time based on a priori data on intensities bounce

$$R(Work^{1,2,3}_{1,3,8})_{t=0} = 0\ ;$$

$$R(Not\_work^{1,2,3}_{1,3,8})_{t=0} = 1;$$

$$R(Work^{2,3}_{IE-CAS,CAS-SPP})_{t=0} = 0\ ; \tag{16}$$

$$R(Not\_work^{2,3}_{IE-CAS,CAS-SPP})_{t=0} = 1\ ;$$

$$R((Work^{1,2,3}_{1,3,8})_t\ /(Work^{1,2,3}_{1,3,8})_{t-1}) = 0{,}1\ ;$$

$$R((Work^{2,3}_{IE\_CAS,CAS\_SPP})_t\ /(Work^{2,3}_{IE\_CAS,CAS-SPP})_{t-1}) = 0{,}1$$

Sets of risk of failures at the current moment of time, taking into account the previous state of subsystems and intersystem communications, can be within the level of risk of failure is estimated as minimal, the consequences of an accident are minimal at:

$$R((Not\_work^{1,2,3}_{1,3,8})_t\ /(Work^{1,2,3}_{1,3,8})_{t-1}) = 0{,}1 - 0{,}2\ ; \tag{17}$$

$$R((Not\_work^{2,3}_{IE\_CAS,CAS\_SPP})_t\ /(Work^{1,3}_{IE\_CAS,CAS-SPP})_{t-1}) = 0{,}1 - 0{,}2\ ;$$

the risk failure level is assessed as acceptable, the consequences of the accident are insignificant at:

$$R((Not\_work^{1,2,3}_{1,3,8})_t\ /(Work^{1,2,3}_{1,3,8})_{t-1}) = 0{,}2 - 0{,}37\ ; \tag{18}$$

$$R((Not\_work^{2,3}_{IE\_CAS,CAS\_SPP})_t\ /(Work^{1,3}_{IE\_CAS,CAS-SPP})_{t-1}) = 0{,}2 - 0{,}37\ ;$$

the risk failure level is estimated as maximum, the consequences of the accident are significant at:

$$R((Not\_work^{1,2,3}_{1,3,8})_t\ /(Work^{1,2,3}_{1,3,8})_{t-1}) = 0{,}37 - 0{,}63\ ; \tag{19}$$

$$R((Not\_work^{2,3}_{IE\_CAS,CAS\_SPP})_t\ /(Work^{2,3}_{IE\_CAS,CAS-SPP})_{t-1}) = 0{,}37 - 0{,}63;$$

the failure risk level is assessed as critical at:

$$R((Not\_work^{1,2,3}_{1,3,8})_t\ /(Work^{1,2,3}_{1,3,8})_{t-1}) = 0{,}63 - 1\ ; \tag{20}$$

$$R((Not\_work^{2,3}_{IE\_CAS,CAS\_SPP})_t\ /(Work^{2,3}_{IE\_CAS,CAS-SPP})_{t-1}) = 0{,}63 - 1.$$

The construction and study of BBN failure risk assessments FS, FC, FIC and FI CTS was carried out using the software product GeNIe [20].

The use of the GeNIe environment makes it possible to diagnose the TC of each FS, FC, FIC and FI CTS. Perform a regression analysis of the influence of network each parent element on its corresponding child element. Implement a graphical display of the failures risk assessment predicting results FS, FC, FIC and FI CTC. Calculate the values of the working capacity loss probability, damage from failures and assessments of failures risk FS, FC, FIC and FI CTC.

When modeling the SPP's BBN (Fig. 2), for various failure risk values the input component, the failures risk values of functionally interconnected and interacting FS for 20,000 hours SPP operation were determined (Fig. 3). The operating state and failure, for example, of the CS subsystem for the risk of failure at the input element of the SPP 0.26 when simulating the BBN of the SPP is shown in Fig. 4.

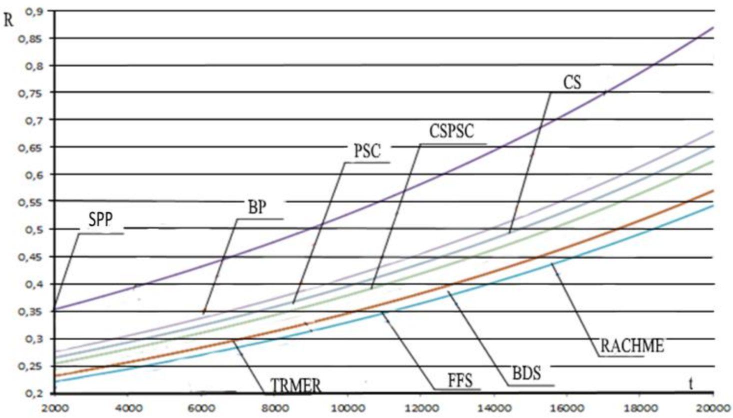


Fig. 3. SPP subsystems predictable risk

The purpose of using BBN in assessing the risk of failures FS, FC CTC is a posteriori conclusion. The a priori data are dynamically recalculated and form a posterior failure risk estimate, which is a priori information, to process the new information. Post hoc inference is based on procedures for analyzing data obtained from the use of BBN. When implementing this approach in research, modeling on a priori and a posteriori data, the predicted TC FS, FC of the power plant are determined, which have the greatest impact on the performance of the main engine and the operation of the entire system for various time periods. It follows from the research results that the predicted maximum non-operating state during the operation of the SPP is 20,000 hours. Corresponds to the most vulnerable subsystems ME, VR (Fig. 2).

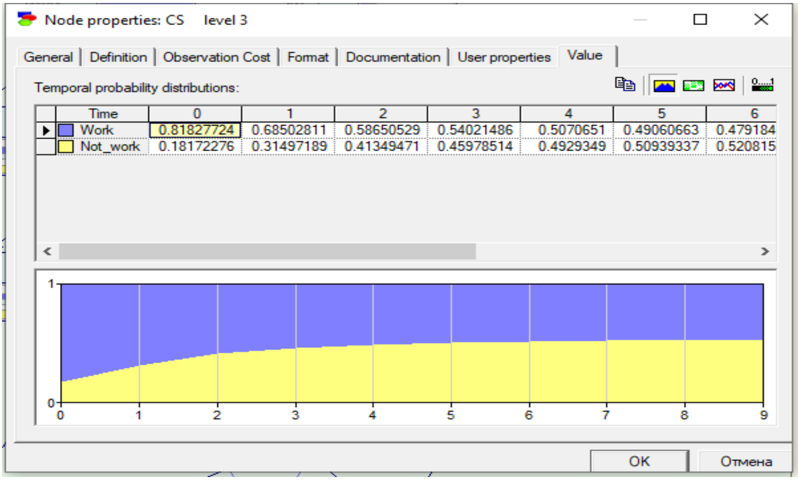


Fig. 4. Operating state and failure of the CS subsystem for failure risk the input component of the SPP 0.26

Because subsystems ME, VR are dependent at the level of the hierarchical structure of the SPP, therefore, in the future, it is necessary to regularly check the subsystems in order to find possible causes of their failure, thereby increasing operation reliability of ME, VR, and hence the whole SPP at all. Thus, based on the intellectualization TC assessment FS, FC, FIC and FI of the CTS by diagnostic features, it is possible to substantiate the forecast failures risk FS, FC, FIC and FI of the SPP.

The considered principle intelligent system functioning, its structure, in terms of the technical and technological foundations of construction on the example of SPP, reflected in the method and model for assessing and predicting FS, FC, FIC and FI CTS failures risk can be considered as conceptual task. The described method, the developed model of an intelligent system for assessing and predicting the risk of CTS failures on network infrastructures, as a research result, confirmed the relative insensitivity to incomplete technological data FS, FC, FIC and FI.

Application of research results allows providing:

– formation of principles for the intelligent system for diagnosing and predicting the CTS failures risk construction and operation;

– intellectualization model of TC estimation and forecasting ship CTS failures risk oby diagnostic features, which has a relative insensitivity to incomplete technological data FS, FC, FIC and FI CTS creation;

– intellectualization model for ES evaluation based on the use of a priori information about failures, linking the types of TC FS, FC, FIC and FI of complex systems and their diagnostic features in the failure risk form creation;

– identifying the most vulnerable FS, FC, FIC and FI CTS and solving the problem of determining the failures causes depending on failures risk in the TC diagnostics.

4. Conclusion

The results of the development of a diagnostic model for a complex technical system with incomplete technological data and its implementation in an intelligent system for assessing and predicting FS, FC, FIC and FI ship CTS failures risk made it possible to obtain a priori information about the technical condition of a complex system. The types of technical condition FS, FC, FIC and FI are determined on the basis of diagnostic features of a complex system using ship power plant example. Predicting complex technical system technical state type was carried out using a posteriori inference in Bayesian belief networks. The conducted studies presented in the article evaluated the results of functioning of an intelligent system for diagnosing and predicting complex technical system failures risk, which makes it possible to identify the most vulnerable FS, FC, FIC and FI CTS and predict their TC.

The model for diagnosing and predicting the risk of failures of subsystems, components, interconnections can be considered as an intelligent system conceptual model for diagnosing and predicting complex technical systems failures risk on network infrastructures, which has a relative insensitivity to incomplete technological data.

The use of the developed method and model, taking into account the hierarchical levels FS, FC, FIC and FI, when searching for the causes of failures in complex technical systems, allows us to control failures risk in systems when information about failures in their structures is received according to TC. The application of the method and model allows predicting trends in the risk of system failures, taking into account changes in individual FS failures risk of FC, FIC and FI in order to further choose a strategy for their restoration or replacement.

5. References

[1] V.Vychuzhanin, N. Rudnichenko, Metody informatsionnykh tekhnologiy v diagnostike sostoyaniya slozhnykh tekhnicheskikh sistem, monografiya, 2019.

[2] V. Vychuzhanin, N. Rudnichenko, Devising a method for the estimation and prediction of technical condition of ship complex systems, Eastern-European Journal of Enterprise Technologies, 84 (2016) 4-11. DOI:10.15587/1729-4061.2016.85605.

[3] V. Vyuzhuzhanin, N. Rudnichenko, N. Shibaeva, Data Control in the Diagnostics and Forecasting the State of Complex Technical Systems, Herald of Advanced Information Technology, 2 (2019) 183-196. DOI:10.15276/hait.03.2019.2.

[4] V. Vychuzhanin, N. Rudnichenko, Complex Technical System Condition Diagnostics and Prediction Computerization, in: CMIS-2020 Computer Modeling and Intelligent Systems / Ceur-ws.org. 2608.2020. pp.1-15. DOI:10.32782/cmis/2608-4

[5] J. Sшrensen, Marine Control Systems Propulsion and Motion Control of Ships and Ocean Structures, 2013.

[6] I. Lazakis, Advanced ship systems condition monitoring for enhanced inspection, maintenance and decision making in ship operations, Transportation Research Procedia, (14) 2016 1679 – 1688. DOI:10.1016/j.trpro.2016.05.133

[7] P. Zhang, Z. GaoCao, L. Dong, Marine Systems and Equipment Prognostics and Health Management. Systematic Review from Health Condition Monitoring to Maintenance Strategy, Machines, 10 (2022) 72-79. DOI:10.3390/machines10020072.

[8] F. Lan, Y. Jiang, Performance Prediction Method of Prognostics and Health Management of Marine Diesel Engine, in: Proceedings of the 2020 3rd International Conference on Applied Mathematics, Modeling and Simulation, Shanghai, China, 20–21 September 2020, 2020 vol 1670. pp.1-8. DOI:10.1088/1742-6596/1670/1/012014.

[9] D. Boullosa-Falces, J. Barrena, Monitoring of fuel oil process of marine diesel engine, Appl. Therm. Eng, 127 (2017) 517–526.

[10] P. Compton, Troubleshooting Marine Diesels, International Marine, 2017. DOI:10.1016/j.applthermaleng.2017.08.036

[11] Z. Jin, L. Meng, J.M. Guerrero, R. Han, Hierarchical Control Design for Shipboard Power System with DC Distribution and Energy Storage aboard Future More-Electric Ships, IEEE Transactions on Industrial Informatics, 14 (2018) 703-714. DOI:10.1109/TII.2017.2772343

[12] V.V. Vychuzhanin, N.D. Rudnichenko, A.V. Vychuzhanin, A.Ye. Kozlov Programnoye obespecheniye avtomatizatsii otsenki riska otkazov slozhnykh tekhnicheskikh system, Informatics and Mathematical Methods in Simulation, 4 (2019) 315-323.

[13] P.C. Lucienne An Introduction to Probabilistic Risk Analysis for Engineered Systems, 2011. DOI:10.1002/9780470400531.eorms0680

[14] R. Wang Chen, C. Guan, A Bayesian inference-based approach for performance prognostics towards uncertainty quantification and its applications on the marine diesel engine, ISA Trans, 118 (2021) 159–173. DOI: 10.1016/j.isatra.2021.02.024
[15] F.V. Jensen, Bayesian Networks and Decision Graphs, 2007.
[16] D. Handayani, W. Sediono. Anomaly Detection in Vessel Tracking: A Bayesian Networks (Bns) Approach, International Journal of Maritime Engineering (RINA Transactions Part A) 157 (2015) 145–152. DOI:10.3940/rina.ijme.2015.a3.316
[17] R. Ferdous, F. Khan, R. Sadiq, P. Amyotte, B. Veitch, Fault and Event Tree Analyses for Process Systems Risk Analysis: Uncertainty Handling Formulations, Risk Analysis 31 1 (2011) 86–107. DOI: 10.1111/j.1539-6924.2010.01475.x
[18] E.C. Harrington, The desirable function. Industrial Quality Control, 2110 (1965) 494-498.
[19] OREDA. OREDA-Offshore Reliability Data Handbook, 2015.
[20] Genie9: Making Local Backup Easy and Affordable - Genie9, 2022. [Online]: https://www.zoolz.com/genie9/.

УДК 004:37:001:62

ІНТЕЛЕКТУАЛІЗАЦІЯ ДІАГНОСТИКИ СКЛАДНИХ ТЕХНІЧНИХ СИСТЕМ

Вичужанін В., Вичужанін А.
Національний університет "Одеська політехніка"
E-mail: v.v.vychuzhanin@op.edu.ua

Анотація. *У статті наведено результати розробки моделі діагностування суднової складної технічної системи з неповними даними та її реалізації в інтелектуальній системі оцінки ризику відмов підсистем, компонентів, міжкомпонентних зв'язків, що дозволяє отримати апріорну інформацію про технічний стан складної системи. Визначено види технічного стану підсистем, компонентів, міжкомпонентних зв'язків на основі діагностичних ознак складної системи на прикладі суднової енергетичної установки для оцінки ризику їх відмов. Прогнозування типу технічного стану складної технічної системи було здійснено за допомогою апостериорного висновку в байєсівських мережах вірувань. Дослідження, представлені в статті, оцінювали ризики відмов у результаті використання інтелектуальної системи діагностики та прогнозування ризиків відмов складної технічної системи судна. Модель діагностики та прогнозування ризику відмов підсистем, компонентів, взаємозв'язків можна розглядати як концептуальну модель інтелектуальної системи діагностики та прогнозування ризику відмов складних технічних систем на мережевих інфраструктурах, яка має відносну нечутливість до неповної технологічні дані.*

Ключові слова: *технічний стан; складна технічна система; ризик невдачі; діагностика; прогнозування; інтелектуальна система; мережа переконань Байєса; нечутливість до неповних даних.*

Section 8. Multi-agent systems and distributed computing

UDC 004.89

THE DEVELOPMENT OF AGENT-BASED LEARNING PLATFORM WITH TEMPORAL LOGIC SPECIFICATIONS

Dr.Sci. Axak N., Ph.D. Kushnaryov M., Tatarnykov A.
Kharkiv National University of Radio Electronics, Ukraine
ORCID: 0000-0001-8372-8432, 0000-0002-3772-3195,
0000-0002-1632-8188
E-mail:nataliia.axak@nure.ua, maksym.kushnarov@nure.ua,
andrii.tatarnykov@nure.ua

Abstract. *A model of an agent-based learning platform (ALP) has been developed, in which, based on the combination of agent technology and computer vision, an approach is proposed that allows combining collective control, coordination and cooperation of agents to provide online educational services. The proposed approach allows, firstly, students to be aware of the progress of their own learning activities, and secondly, to notify parents in the event of a negative incident, i.e., when their child does not attend classes according to the schedule established in this educational institution, or this student received a score based on the results of the final/semester academic performance check in accordance with the schedule of the educational process less than the minimum allowable. The provision of such services is realized through the student behavior monitoring subsystem and the community of intelligent agents.The formal specification of agents was made using alternating-time temporal logic (ATL). The formalization is used to model the simultaneous behavior of a system with an unlimited number of agents and states, as well as to verify some properties expressed as ATL formulas.A software prototype of an agent environment for integrating distributed agents into a single information space that combines the Moodle learning platform, a university web portal, a Student behavior monitoring subsystem developed and Amazon S3 cloud storage was developed*
Keywords: *learning management system, monitoring, Moodle, learning process, agents, multiagent system, temporal logic specifications, alternating-time temporal logic*

1. An introduction to achievements of multi-agent systems in education

The development of computing systems and the Internet of Things (IoT) has led to the ubiquitous distribution of distributed computing, embodying the successful interaction of disparate objects. In this regard, we see the widespread distribution of sets of autonomous computing objects (agents), called multi-agent systems (MAS). It is thanks to the effective application of agent-oriented technologies that the development of distributed and intelligent programs in complex and highly dynamic environments became possible. Multi-agent systems are successfully implemented in almost all fields of activity, including production, education, marketing, health care, etc., improving people's daily life. Getting a good education has always been an important aspect in our lives. Most citizens consider a good education a means of obtaining a decent life. The globalization of the world and the globalization of learning are expanding the application areas of artificial intelligence (AI) in education, including profiling and prediction, assessment, personalization, and intelligent learning systems. The main goal of educational institutions is to improve the quality of education, that is, to improve the process of acquiring knowledge by students.In addition, we observe an approach to the actual problem of higher education institutions, which are forced to close due to quarantine or due to the war that Russia has unleashed in Ukraine, and switch to online education using virtual environments. The main problem facing learning in virtual environments is not only the availability of educational content for users, but also the ability to present knowledge in the right place, at the right time, and in the right way. For this reason, researchers are paying more and more attention to the application of intelligent agents to manage the learning process in order to improve the performance of students in the online environment. Therefore, the increasing integration of intelligent agents into our lives, especially where computer systems must be able to communicate both with each other and with people, requires research and development of new concepts and tools in the complex field of multi-agent systems.

2. Related works

The technology of multi-agent systems is used in almost all real-world applications, whether it is a simple e-commerce auction or an air traffic control system. It is one of the new technologies that has gained popularity very quickly due to the fact that it allows easy development of complex and distributed systems. It also appeared in the educational environment. Several multi-agent e-learning systems containing different functions have been developed and implemented. Paper [1] analyzes and compares various existing multi-agent e-learning systems based on their characteristics such as interactivity, adaptability, security, etc. The purpose of the special issue “Advances in Multi-Agent Systems” is to promote MAS by increasing its visibility and increasing its accessibility to the scientific community [2].

This review aims to provide an overview of current research on MAS agents and technology and highlight the high level of activity in this area. The researched works demonstrate the constant growth of the scientific community's interest in new models, techniques and methods for multi-agent systems. Although research related to MAS is still developing in the most "classical" directions, at the same time it is also expanding into new areas - learning huge amounts of data. The purpose of another special issue "New Insights in Multi-Agent Systems Cooperation, Control and Optimization" [3] is to attract new, high-quality contributions to the theory, modeling, development and application of social processes of multi-agent systems. Approaching many problems in science and technology as multi-agent systems emphasizes the need to investigate problems related to collective control, collective decision-making and optimization, collective and social machine learning, multi-agent reinforcement, as well as coordination, cooperation and evolution of agents, etc. The use of agents in learning systems is considered in [4, 5, 6, 7, 8, 9, 10, 11]. The purpose of the work [12] is to analyze the use of software architectures in intelligent e-learning systems. In addition, the authors propose an open, distributed, agent-based software architecture for flexible, personalized learning and the development of personalized educational resources. They also discuss an ontology-based information model for personalized learning as part of a proposed agent-based software architecture. The relationships between the components of the presented software model and the information model based on the ontology, which describes the data and knowledge necessary for the effective operation of the learning system, are also considered. The document [13] presents the architecture of the intelligent learning management system (ILMS) applied to Moodle. They presented the design and implementation of an agent that selects a learning strategy according to a student's learning style. The selected learning strategy used to filter the learning objects displayed to students. The authors of the work [14] believe that in order to achieve the main goal of electronic learning systems (LMS) - increasing the effectiveness of learning, it is necessary to satisfy individual requirements and provide personalized practices depending on the abilities of students. This paper proposes a dynamic multi-agent system that includes five agents that takes into account the differences in capabilities of different users: the project clustering agent is used to cluster a set of educational resources/projects into similar groups; the student cluster agent groups students according to their preferences and abilities; a student and project matching agent is used to map each student group to an appropriate project or specific learning resources according to specific design criteria; the student matching agent is designed to efficiently match different students; dynamic student clustering agent is used for continuous tracking and analysis of student behavior in the system, such as changes in knowledge and skill levels.

The article [15] presents a multi-agent system for recommending learning objects in virtual learning environments aimed at improving the customization of learning instructions regarding learning content according to the student's profile. The structure of the system consists of four intelligent agents: administrator agent (responsible for student access control); profile agent (responsible for determining the learning style of students); recommendation agent (analyzes student history to recommend educational facilities); tutoring agent (offers help to students). The document [16] proposes an approach to building a personalized e-learning environment, in which the main emphasis is placed on the development of students' needs. An adaptive agent-based architecture is proposed that extends the Moodle platform to support learning solutions and behavior adaptation. The paper describes the characteristics, functions, and interactions of the agents involved in each module of the adaptive architecture, as well as the intelligent agent for making instructional decisions. The purpose of this agent is to gather information generated by other agents and provide the best personalized support for end users, teachers and students, taking into account their relationship to the learning environment. The aim of the article [17] is to present the benefits of integrating ubiquitous computing together with distributed artificial intelligence techniques to build an adaptive and personalized context-aware learning system using mobile devices. The authors proposed a multi-agent context-dependent e-learning system with the following functionalities: context-dependent learning planning; personalized course evaluation;

selection of educational facilities according to the student's profile; search for educational objects in repositories; search for thematic teaching assistants; access to current context-oriented learning activities. The system [18] is focused on three main characteristics: the learning style according to the Felder-Silverman learning style model, the level of knowledge and possible defects of the student. Three types of disabilities were taken into account, namely hearing impairment, visual impairment and dyslexia. The system will be able to provide students with a sequence of learning objects that matches their profiles for personalized learning. The main goal of this system is to recommend to students a learning path that matches their characteristics and preferences using the Q-learning algorithm. The main goal of the study [19] is to build a multi-agent architecture that provides the ability to adapt to the student, preferences in the e-learning environment, analyze and control communication and interaction between various agents of the proposed system. To this end, they introduced a distributed intelligent whiteboard agent that provides communication between participating agents. The document [20] proposes a multi-agent architecture of an emotional-intellectual e-learning system that aims to help children with autism spectrum disorder overcome learning disabilities. The proposed architecture is based on several agents that allow to intelligently solve emotional, cognitive and pedagogical problems. The document [21] describes an agent-oriented approach, which is aimed at creating learning situations by solving problems. The proposed system is designed as a society of agents that organizes interfaces, coordinators, information sources and mobile devices. The goal of this approach is to challenge learners to engage in multiple learning activities chosen according to their skill level and preferences to enable adaptive learning and reduce the number of learners in an electronic environment. The aim of the study [22] is to develop an e-learning system for basic mathematics that is able to provide each student with personalized content to overcome misconceptions. The system uses a multi-agent architecture to monitor student activity while simultaneously observing and modeling student knowledge and misconceptions. Lessons and exam questions are selected dynamically by the multi-agent system to cover the prerequisites of new lessons depending on the user's profile. In order to provide intelligent tuning and improve the adaptation quality of the learning environment, [23] proposes a multi-agent adaptive learning system architecture based on incremental hybrid case-based reasoning to make the right decision in real time under the current learning environment. The documents [24, 25, 26, 27] presents mobile learning systems using the multi-agent paradigm. The systems provides adaptive learning content, personalized to the learner's style and preferences, to increase learner satisfaction and facilitate the learning process. Taking into account the research works reviewed earlier, it is found that agents can provide students with personal assistants for personalized learning in order to improve the effectiveness of the learning process. However, the problems that arise during distance learning, namely the monitoring of class attendance and the detection of falsification during testing or exams, still remain unsolved. The solution to such an important scientific and practical problem as the development of an agent-based learning platform can be achieved through the introduction of modern methods and concepts into the educational process in the following ways. Personalized education. The development of the platform should be aimed at individualising learning, taking into account the needs and capabilities of each student. The use of an agent-based approach will allow creating agents for each student that will take into account their personality, skills, learning pace, etc. Use of artificial intelligence to analyze progress. Intelligent agents can use data analytics and machine learning to assess learner progress, identify problem areas, and recommend further training. This can lead to more effective learning and improved outcomes.

1. Promoting cooperation and interaction. The design of the platform should allow for collaboration between students and teachers, sharing of experiences and knowledge. Agents can coordinate this process and encourage mutual assistance.
2. Monitoring and prevention of social problems. Agents can serve as a monitoring system to identify students who are at risk from social aspects. This can include timely detection of alarming situations and alerting parents or educators.

3. Access to education for all. Ensuring accessibility and equal opportunities for all students. The development of the platform should be focused on providing education to students of different categories and social groups.

Integration of these aspects into the development of the platform will allow raising important scientific and practical problems in the educational sphere and contribute to their further solution through the use of modern technologies and methods. Statement of the problem - developing the prototype of an agent-based learning platform (ALP) to provide online education services, which allows students to be aware of the progress of their own educational activities and to interact with parents in the event of an incident when their child is at risk.

To solve this problem, you need to take the following steps:

1. Define the requirements and functionality of the system:

– the ability to track student progress;

– mechanisms of interaction between students and parents, including notifications and consultations.

2. Designing the system architecture:

– consideration of an agent-based approach and identification of agents (students, teachers, parents);

– defining the system architecture, including modules for interaction and communication between agents.

3. Selection of technologies and stack:

– selection of programming languages, frameworks and technologies to implement the platform's functionality.

4. Development of prototype functionality:

– creation of modules for registration of students, teachers and parents;

– implementation of progress tracking and learning assessment capabilities.

5. Implementation of interaction between agents: development of mechanisms for communication and data exchange between students, teachers and parents; - creating a notification system to report student progress and possible incidents.

6. Testing and debugging:

– testing the system depending on the requirements and functionality;
– identify and eliminate errors and deficiencies.

3. The proposed model of the agent-based learning platform to provide online education services

Learning management processes become efficient and transparent when each participant in the educational process will have access to the information intended for him. This is possible when internal management processes at the university are carried out with the help of information technologies. Educational services become understandable and accessible in electronic form, and the management of the educational institution will always have true data to make effective decisions. For effective management of internal learning processes in an educational institution, the authors have developed a model of an agent-oriented educational platform, which can be used to provide services to university management, teachers, students, and their parents. The proposed ALP model is a general structure that covers the relationship of all components: a multi-agent system (MAS), a university web portal, a Moodle learning management system, a cloud data storage and a developed student behavior monitoring subsystem [28]. The architecture of the proposed agent-oriented learning platform for providing online educational services is shown in Figure 1.

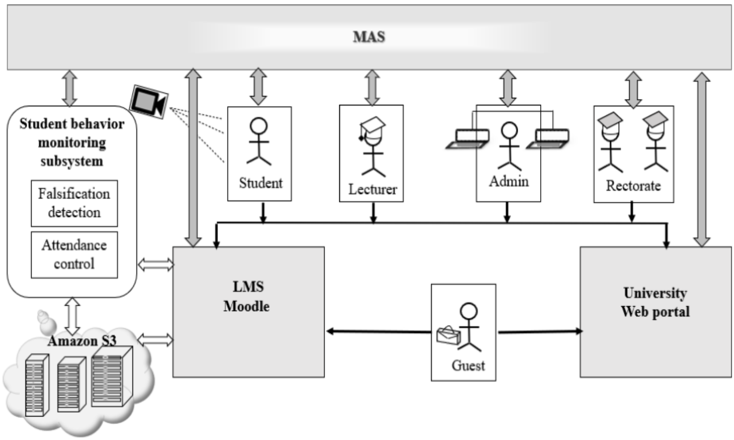


Fig. 1. Architecture of the agent-oriented learning platform

In Figure 1, user access to ALP components is indicated by black arrows; the connection of ALP components for monitoring student behavior is indicated by white arrows; the messaging relationship is indicated by the gray arrows. The multi-agent system is designed for modeling the process of communication and coordination of participants in the educational process. This approach helps to solve urgent problems of each of the participants thanks to the monitoring of information located on Moodle and on the university's Web portal.

The multi-agent system is presented as follows

$$MAS = \{Ag, A, E\}, \quad (1)$$

where $Ag = \{Ag_{St}, Ag_{Lc}, Ag_{Ad}, Ag_{Rc}, Ag_{Gs}\}$ – a set of agents operating in the environment E; $A = \{A_{St}, A_{Lc}, A_{Ad}, A_{Rc}, A_{Gs}\}$ – a set of agent actions ($A_i = \{a_1^i, a_2^i, \ldots, a_n^i\}, i = \overline{(Student, Lecturer, Admin, Rectorate, Guest)}$) ; $E = \{E_{Monitoring}, E_{Moodle}, E_{University}\}$ – a set of states of the environment of Student behavior monitoring subsystem ($E_{Monitoring}$), the environment of LMS Moodle (E_{Moodle}) and the environment of University Web portal ($E_{University}$),
($E_j = \{e_1^j, e_2^j, \ldots, e_m^j\}, j = \overline{(Monitoring, Moodle, University)}$). For convenience, we use abbreviations: instead of $Student$ we will use $St, Admin \rightarrow Ad$, $Rectorate \rightarrow Rc$, $Guest \rightarrow Gs$, $Monitoring \rightarrow Mn$, $Moodle \rightarrow Md$, $University \rightarrow Un$.

Each agent is assigned a specific role. MAS functionality is available to agents with the following roles:

Ag_{St}, – students, graduate students and doctoral students;

Ag_{Lc} – lecturers and researchers;

Ag_{Ad} – site administrator;

Ag_{Rc}– university administration (rector, vice-rectors, dean, etc.);

Ag_{Gs} only if the role of the guest is performed by the student's parents.

The interaction of agents and the environment is the execution of actions:

$$f_j : e_0^j \xrightarrow{a_0^i} e_1^j \xrightarrow{a_1^i} e_2^j \xrightarrow{a_2^i} \ldots \xrightarrow{a_{u-1}^i} e_u^j \xrightarrow{a_u^i} \ldots \quad (2)$$

The behavior of the environment is modeled as a function of the state converter

$$\Upsilon_j : f_j\left(ended\ with\ a_k^i\right) \rightarrow \rho(E),\ (k = \overline{1, n}). \quad (3)$$

This function represents the set of environmental states $\Upsilon_j(f_j)$ that can occur as a result of action a_n^i in state e_m^j. Due to the fact that learning resources are decentralized, the use of only one stationary agent will not be sufficient. Therefore, mobile agents are needed to receive information from various resources (LMS Moodle, University Web portal, Amazon S3). A society of agents was created to develop a multi-agent system (Fig. 2).

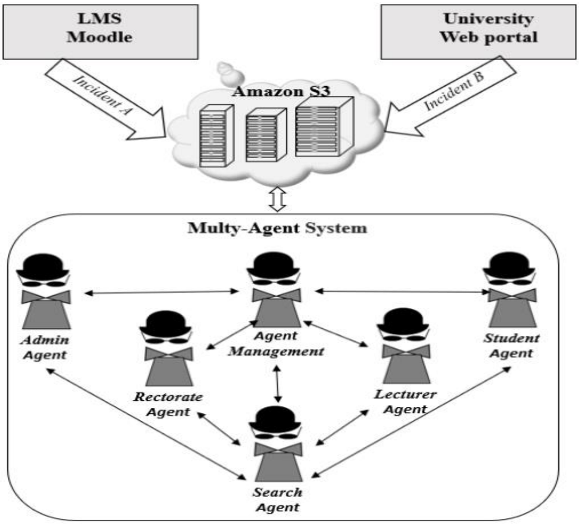


Fig. 2. The main actors of the multi-agent system

Student behavior monitoring subsystem allows: the lecturer to make a decision regarding the assessment of the student's knowledge in case of detection of falsifications during testing or exams; the lecturer to increase the student's activity by adjusting the total score for attending classes; determine the duration of students' attendance at classes in accordance with the established schedule in the educational institution for parental control.

All users of the proposed ALP - employees, admin, students, postgraduates, applicants, guests, etc., can have access to various components of the ALP depending on the role performed and receive information and all kinds of help regarding the educational process at the institution of higher education. For example, agent Ag_{St}, can:

– have access to Moodle blocks related to store, communicate and collaborate;

– automatically receive up-to-date information from the university's Web portal, from the dean's office and lecturers;

– participate in surveys (internal and external);

– learn about scientific conferences, internships, scientific circles, Olympiads, hackathons, etc.

Agent Ag_{Lc} can: have access to Moodle blocks related to store, communicate, evaluate and collaborate; to receive information about the duration of students' attendance at classes and notification of detection of falsifications during testing or exams; automatically receive up-to-date information from the university's Web portal, from the dean's office and lecturers.

Agent Ag_{Ad} can: performs the necessary actions to support the operation of the Moodle platform; receives up-to-date information for the effective operation of the Moodle platform and the university's Web portal; is responsible for maintaining the functionality of the university's Web portal and ensuring network security, manages the placement, updating, and moderation of content.

Agent Ag_{Rc} can: have access to Moodle blocks; provide information for the university's Web portal.

Agent Ag_{Gs} can: have access to open information on the university's Web portal; receive information about the implementation of the student's individual study plan, if the role of "guest" is performed by parents.

Agents are represented as functions

$$Ag: f_j\left(ended\ with\ e_u^j\right) \to A, \ \ (u = \overline{1, m}). \tag{4}$$

The behavior of an agent in an environment is the set of all runs

$$F(Ag, Env), \qquad (Env = \{E, e_0^j, Y_j\}). \tag{5}$$

Moodle LMS is a learning platform used to provide lecturers and students with a personalized learning environment. This software is deployed on its own web server. Moodle is free and open source software under the GNU General Public License. Anyone can adapt, extend or modify Moodle for both commercial and non-commercial projects without any licensing fees. There are many tasks in the educational process. At the initial stage of the development of an agent-oriented educational platform, the authors limited themselves to the implementation of only one task - the assessment of the knowledge acquired by the student. This functionality provides important information related to a student's progress and allows you to individually highlight students' weaknesses or strengths and inform parents of their child's academic progress. In order to achieve this goal, there is attendance monitoring and ongoing control, which is carried out during the semester during classroom classes (lectures, laboratory and practical classes, etc.) To do this, according to the established schedule of classes, "Student Agent" detects negative aspects: absence of a student in classes; sum of points obtained according to the results of the final/semester test of success in accordance with the schedule of the educational process is less than the minimum permissible. In case of negative aspects, "Student Agent" sends a message to parents that their child is in the risk zone. Figure 3 shows a usage diagram showing the main functionality of the agents. Initialization occurs as the first event handled by the agent, and deinitialization occurs after the application terminates. Deinitialization and initialization are performed only once. Receiving messages from other agents is one of the main functions of an agent. An agent can receive messages from other agents or information from the network. Messages are sent to inform other agents or to give an indication of what action to take.

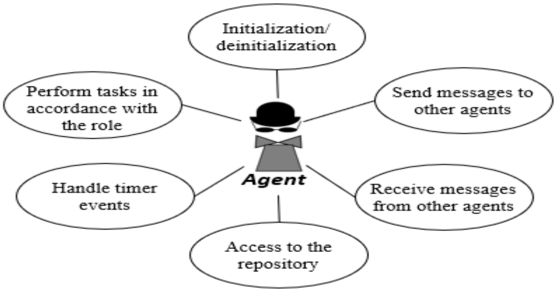


Fig. 3. Agent program usage chart

Receiving messages from other agents is one of the main functions of an agent. An agent can receive messages from other agents or information from the network. Messages are sent to inform other agents or to give an indication of what action to take. Timer event handling is used to periodically check any resource on the network or to check the status of any task performed by another agent. Each agent is equipped with a knowledge base consisting of scenarios to follow in order to achieve a given goal.

A multi-agent system is based on several agents that are designed to communicate and exchange data in order to provide users with relevant information, for example, to notify a student and his parents that he is at risk due to poor academic performance. Within the framework of the proposed model, the agent receives messages from other agents and external notifications about the activity of the learning process from the environment.

Agent Management provides agent management, allows you to create and delete agents, publish information about the services provided and find agents providing the necessary services and enter into negotiations with them. The diagram of the interaction performed by each agent from the moment of launch to the end of its existence in the environment is presented in fig. 4.

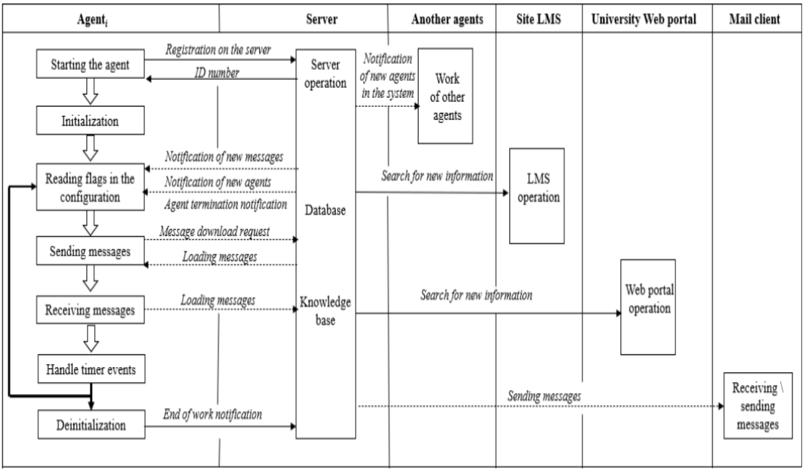


Fig. 4. The agent activity (cooperation) diagram

Actions performed by participants, who are involved in ALP, are marked with rectangles. The sequence of actions performed by agents is indicated by vertical arrows, and messages exchanged by agents are shown by horizontal arrows. A solid arrow indicates that this message is sent in any case, and a dashed one means that sending the message depends on some condition. Viewing the chart starts with the agent running. At this point, a request to register on the server is broadcast. The server, upon receiving such a request, adds information about this agent - the current location and its type (Agent Management, Admin Agent, Rectorate Agent, Lecturer Agent, Search Agent and Student Agent) to its database. Based on this data, it generates a unique identification code for each agent. This code is used when sending messages between agents and the server. After registering on the server, the agent executes a script specific to it. Messages are received and sent if the corresponding script is executed for this agent. The agent sends a request to receive data, the server sends it to the agent, and then the script is launched (Fig. 5). The downloaded message is stored in a file on the drive, and its handle is passed to the script.

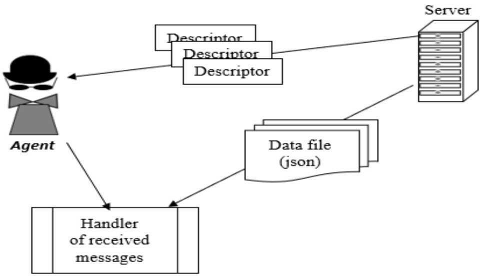


Fig. 5. Messaging mechanism

When sending a message, the corresponding script is launched, which is designed to process the data that will be sent. Then a descriptor is formed, which specifies the size of the message, the way to place the message on the disk, the identifiers of the agents that should receive it, and a checksum to verify the integrity of the messageThis handle is sent to the server, which downloads this message at the specified address. It then parses the handle and sends a new message notification to all agents specified in the handle.

Thus, the developed system gives ALP users the opportunity to maintain updated information that helps to improve learning outcomes.

3.1. Possible Action Specifications of Agents

Most formal specification methods are based on some variant of temporal logic. The most widely used is Computation Tree Logic (CTL). Its model of time is a tree structure in which the future is not defined; there are different paths in the future, each of which can be an actual path that is realised. CTL extends the classical logic of statements with temporal operators. When describing formulas, we use the functional symbols of standard arithmetic operations, the predicate of equality and arithmetic comparison, constants corresponding to each of the objects of the subject set, and the logical constants true and false. We generate formulas according to the Hennessy-Milner Logic rules. [29]. Let *MAS* be the implementation of a multiagent system, using the syntax of micro Common Representation Language 2 (mCRL2). mCRL2 is based on the algebra of communication processes, in which the fundamental concept is the process. Each process has a corresponding state space, or labelled transition system (LTS), which contains all the states that the process can reach and the possible transitions between these states. To check the equivalence of the developed system, it is necessary to determine whether $MAS \equiv S$, where the symbol $\equiv$ is some behavioral equivalence, S is a specification of the expected behavior of the system in mCRL2. To verify the developed model, it is necessary to determine whether $MAS \models F$, where the symbol $\models$ is a satisfaction relation, F defines a desired property of the system. Atomic propositions can be used in logic to find a set of states that satisfies these atomic propositions. The operators $\langle\ \rangle$ and [ ] can be used with a set of actions.

Let's define the modal and temporal properties of the system.

Modal properties describe the capabilities of agents and their limitations in interacting with the environment and other agents. These properties determine the ways in which agents perceive, process, and interact with information from their environment. This means that an agent can have several different states that reflect its current behavior, activity, or response to events.

Some of the modal properties of agents include:

Modes of information perception: An agent can be in different modes of information perception, such as active, passive, or standby.

Levels of responsibility: An agent can have different levels of responsibility, from autonomous agents that make decisions on their own to agents that work within a defined context or under directives.

Levels of skill and experience: An agent can acquire skills and experience over time, which can affect its ability to adapt and solve problems.

Levels of task importance: an agent may place different weights on different tasks and goals, which affects its priorities and focus of action.

Temporal properties indicate aspects of time and the temporal nature of agent behavior. These properties relate to how agents perceive, process, and respond to events that occur at a particular point in time or over a particular time span. Understanding temporal aspects allows for better modelling and analysis of temporal aspects in systems with agents.

Some of the temporal properties of agents include:

Synchronicity or asynchronicity determines whether agents perform their actions and perceive information synchronously (simultaneously) or asynchronously (independently).

Action time constraints indicate time limits for certain actions of agents, for example, the maximum response time to a request, a time limit for completing a certain task, etc.

Timestamps and timers are used to set and track specific points in time, as well as to trigger certain agent actions at a specific point in time or after a specific interval.

Time intervals and frequency indicate how often agents interact, perform certain actions, or observe the environment at certain time intervals.

Time flow modelling takes into account the ability to model and analyze the passage of time in a system of agents, where some actions may be important depending on the time of their execution.

Understanding and correct implementation of temporal properties is important for agent-based systems where time plays a key role.

We will use the concept a labeled transition system (LTS) and process. LTS is a triple $\text{LTS} = (S, A, \rightarrow)$, where transition relation $\rightarrow \subseteq S \times A \times S$.

We use the notation $\left(s \overset{a}{\rightarrow} s'\right)$ to denote that (s,a,s’)∈→.

Formulas F can be generated according to the following rules:

$$F, G ::= false \mid true \mid F \wedge G \mid F \vee G \mid \langle a_n^i \rangle F \mid [a_n^i] F, \quad (6)$$

where a_n^i ranges over A_i ($n \in \mathbb{N}$).

Interpretation of the formula (6):

F may be $true$ in each state or $false$ in every state.

$F \wedge G$ means $true$ in state s only if F and G in a state of s.

$F \vee G$ may be $true$ in state s, if one of the F or G in a state of s.

$\langle a_n^i \rangle F$ means that there is a choice to perform an action a_n^i and reach a state in which F satisfied, $\langle A_i \rangle F ::= \langle a_1^i \rangle F \vee ... \vee \langle a_n^i \rangle F$.

$[a_n^i] F$ means that no matter how we perform an action a_n^i, we will find ourselves in a state where F, $[A_i] F ::= \langle a_1^i \rangle F \wedge ... \wedge \langle a_n^i \rangle F$, (in particular $\langle \emptyset \rangle F = false$ r and $[\emptyset] F = true$).

Using this semantics, we can map formulas F to the set of states where the formula is true [30]. The mapping of a formula F to a set of states is defined as ⟦F⟧. The mapping $[\![F \subseteq S]\!]$ is defined inductively as

$[\![true]\!] = S$; $[\![false]\!] = \emptyset$; $[\![F \wedge G]\!] = [\![F]\!] \cap [\![G]\!]$; $[\![F \vee G]\!] = [\![F]\!] \cup [\![G]\!]$; $[\![\langle a_n^i \rangle F]\!] = \langle . a_n^i . \rangle [\![F]\!]$; $[\![[a_n^i] F]\!] = [. a_n^i .] [\![F]\!]$; where $\langle . a_n^i . \rangle S' = \left\{ s \in S \mid \exists s' . s \xrightarrow{a} s' and\ s' \in S' \right\}$, $[. a_n^i .] S' = \left\{ s \in S \mid \forall s' . s \xrightarrow{a} s' \Rightarrow s' \in S' \right\}$.

The program for monitoring student progress using agents in distance learning is implemented in cycles. In each cycle, each agent receives new data from the environment, starts its behavior, i.e. acts. Each agent of the system has its own set of actions. Each action has a corresponding local impact on the agent and a global impact on the system when this action is performed. The global state is the complete state of the system, which consists of the individual states of all agents in the system.

We will use the synchronous communication model to specify the actions of agents. Agent communication has the form $A = \{Ag_i, Ag_j\}$, i.e. agent Ag_i acts with an agent Ag_j, e.g. action *send* (Ag_{Lc}, Ag_{St}, *message*) means that the lecturer's agent Ag_{Lc} sends a message to the student's agent Ag_{St} with the message as an argument. If the student agent Ag_{St} the lecturer's agent wants to answer Ag_{Lc}, it uses the same template. The local states of agents are updated sequentially based on the specification of their actions. All agents can only see the results of the actions in which they participate. This means that for two agents, the history of all communications is known to both of them, so they have the same local view of the system state.

Student agent Ag_{St} can perform the following actions:

– read information $read_{St}$;

– send messages $send_{St}$;

– receive notifications $recv_{St}$;

– do nothing $noop_{St}$.

Lecturer's agent Ag_{Lc} can perform the following actions:

– analyze data on student performance$anal_{Lc}$;

– generate recommendations for improving teaching materials and methods $rMethod_{Lc}$;

– generate recommendations for assessing student performance$rGrade_{Lc}$;

– read information $read_{Lc}$;

– send messages $send_{Lc}$;

– receive notifications; $recv_{Lc}$

– do nothing $noop_{Lc}$.

Agent administratorAg_{Ad} can perform the following actions:

– read information$read_{Ad}$;

– send messages$send_{Ad}$;

– receive notifications $recv_{Ad}$;

– do nothing $noop_{Ad}$.

Agent Ag_{Rc} мо perform the following actions:

– generate constituent documents$rDocum_{Rc}$;

– send messages $send_{Rc}$;

– receive notifications$recv_{Rc}$;

– do nothing $noop_{Rc}$.

Agent Ag_{Gs} can perform the following actions:

– receive notifications $recv_{Gs}$;

– do nothing $noop_{Gs}$.

Each state can define transitions to other states within the same parameter, as well as actions to be performed if the parameter is in this state.

Performing certain actions may generate new obligations between interacting agents. All agents can only see the results of actions in which they participate. This means that for two agents, the history of all communications is known to both of them, so they have the same local view of the system state. To detect negative aspects (student absence from classes of more than 20%, or a student received a score based on the results of the final/semester academic performance check according to the schedule of the educational process that is less than the minimum allowable), the following environmental variables are used for the *m*-th discipline. e_1^{Mn} is interpreted as the time students attend classes in minutes, which takes values in the set {0,1,2,...,100}; *minAtt* indicates that the current absence from classes is less than 10% (takes the value 0, otherwise ¬ 1); e_1^{Md} is interpreted as a current assessment of students' knowledge, which takes values in the set {0,1,2,...,100}; *minPnt* indicates that at the moment the sum of points based on the results of the final/semester academic performance check in accordance with the schedule of the educational process is less than the minimum allowable (takes the value 0, otherwise ¬ 1); e_1^{Un} is interpreted as 1 if a lecture is scheduled, 2 - laboratory/practical classes, 3 - exam/test, 0 - no classes. The fact that a student is in the risk zone is denoted by the variable *danger*, which takes values in the set {0, 1}. The fact of changing the learning path is denoted by the variable *changeWay,* which takes values in the set {0, 1}. The family of variables pr_{ij} (0<i, j<100) is used to describe the level of students' knowledge, which takes the values $low, medium$ or $high$, depending on the values of the variables e_1^{Mn} and e_1^{Md} . The main limitation of CTL is the lack of explicit means of specifying systems consisting of several interacting entities. This limitation is overcome in Alternating-time temporal logic (ATL), which allows explicitly taking into account the multicomponent nature of the system. ATL is a branching time temporal logic that extends the logic of computation trees (CTL) for multiple players. ATL naturally describes the computation in multi-agent systems and multiplayer video games. Quantification in ATL is performed along program paths that are possible achievements in the game.

ATL uses time-varying formulas to create model checking tools to address issues such as susceptibility, validity, and controllability. So, additionally, we will use the following ATL syntax and semantics [31].

Let Π be a set of atomic propositions, and Ag a set of k agents. The set of ATL formulas is inductively defined as follows:

– p, for each $p \in \Pi$.

– $\neg$ F, F $\vee$ G, F $\wedge$ G, F $\rightarrow$ G, where F, G are ATL formulas.

– $\langle\langle Ag_i \rangle\rangle \circ F$, $\langle\langle Ag_i \rangle\rangle \square F$, $\langle\langle Ag_i \rangle\rangle F\, \mathcal{U}\, G$, where F, G are ATL formulas and $Ag_i \subseteq Ag$.

The operator $\langle\langle\ \rangle\rangle$ is a path quantifier; $\circ$ (next), $\square$ (box) and $\mathcal{U}$ (until) are temporal operators. Thus, we will use CTL-equivalent logic that constrains Ag_i to $\emptyset$ or Ag. Additionally, we will introduce the following operators: *belief* F – is an operator used to describe the agent's representations; *desire* F – an operator used to describe the agent's desires. The semantics of this operator is specified with respect to the structure of the agent's desires; *intend* F – is an operator used to describe the agent's intentions. The semantics of this operator is specified by the possible worlds defined by the perception see, beliefs belief and plan of the agent. Let's enter the symbol *progress*, which corresponds to the display of the form $\{1,2,\dots,100\} \times \{1,2,\dots,100\} \rightarrow \{low, medium, high\}$. In our example, the atomic offers of synchronous messaging between agents will be $\{send_i\ notification, recv_j\ notification\}$, $\{send_i\ progress, recv_j\ progress\}$. $i, j = \overline{St, Lc, Ad, Rc, Gs}$.

Actions can create obligations between agents that take the following form $Commitments\ (Ag_i, Ag_j, a_n^i)$, This means that the agent Ag_i obliged to the agent Ag_j perform an action a_n^i. Or there may be a form with contingent liabilities in the form of $condCommitments\ (Ag_i, Ag_j, cond, a_n^i)$, which means that the agent Ag_i obliged to the agent Ag_j perform an action a_n^i only on condition of $cond$. The formalization of the laws of the external environment will be as follows. If there are classes on the schedule, the student's agent Ag_{St} reads data on the student's attendance at these classes and their academic performance

$$\bigwedge_{1\le i\le 3}\left(e_1^{Un}=i\Longrightarrow[read_{St}]\left(\bigwedge_{1\le j\le 100}e_1^{Md}\right)\wedge\left(\bigwedge_{1\le k\le 100}e_1^{Mn}\right)\right)\qquad(7)$$

Assertion (8) describes the patterns of student learning progress. That is, if a student performs all control measures in a timely manner and attends almost all lectures and laboratory/practical classes, this does not lead to the fact that he or she is at risk

$$\wedge_{1\le i\le 3}\left(e_1^{Un}=i\Rightarrow(\wedge_{1\le j\le 100}e_1^{Md}>minPnt)\wedge(\wedge_{1\le k\le 100}e_1^{Mn}\gg minAtt)\right)\Rightarrow\neg\,danger\qquad(8)$$

The discovery of this fact leads to the fact that in this case the student's agent Ag_{St} does nothing $progress(e_1^{Md},e_1^{Mn})=medium\Rightarrow[\,noop_{St}]$.

Statement (9) describes the consequences of not completing the study plan on time. It states that because the student has not completed the control measures in a timely manner and has hardly attended classes, the student is at risk

$$\wedge_{1\le i\le 3}\left(e_1^{Un}=i\Rightarrow(\wedge_{1\le j\le 100}e_1^{Md}<minPnt)\wedge(\wedge_{1\le k\le 100}e_1^{Mn}<minAtt)\right)\Rightarrow[true]\,danger\qquad(9)$$

In this case, the student's agent Ag_{St} notifies the teacher′s agent Ag_{Lc}, that the student is at risk, and if the student does not change his or her attitude to study (elimination of debts), the teacher's agent Ag_{Lc} informs parents of the situation (agent Ag_{Gs}) of such a student

$$progress(e_1^{Md},e_1^{Mn})=low\Rightarrow\begin{bmatrix}Commitments\,(Ag_{St},Ag_{Lc},send_{St}\,progress)\,\wedge\\condCommitments\,(Ag_{Lc},Ag_{Gs},\neg changeWay,send_{Lc}\,progress)\end{bmatrix}$$

Early completion of all control measures with high grades leads to a proposal to change the learning path (10)

$$(\wedge_{1\le i\le 100}e_1^{Md}\ge 95)\Longrightarrow[true]\,changeWay.\qquad(10)$$

In this case, the teacher's agent Ag_{Lc} Sends an offer to the student to change the study trajectory, for example, to join a research activity

$$progress(e_1^{Md},e_1^{Mn})=high\Rightarrow[Commitments\,(Ag_{Lc},Ag_{St},send_{Lc}\,notification)]$$

Absence from scheduled classes does not affect students' academic performance (11)

$$e_1^{Un} = 0 \Rightarrow \neg minPnt. \quad (11)$$

The following properties limit the response time of the system. Property (12) requires that at the end of the lesson, the student's agent Ag_{St} read the data on the student's attendance at these classes and their academic performance

$$\langle\langle\{Ag_{St}\}\rangle\rangle A\, \mathcal{F}(e_1^{Un} \neq 0) \Rightarrow \langle\langle\{Ag_{St}\}\rangle\rangle A\, G_{<90} \Rightarrow [read_{St}]. \quad (12)$$

Statement (13) describes the mental state of the agents.

$$\Lambda_{pr\in\{low,medium,high\}}(progress(e_1^{Md}, e_1^{Mn}) = pr \Leftrightarrow \langle\langle Ag_{Lc}, Ag_{St}\rangle\rangle\ belief\ \ progress(e_1^{Md}, e_1^{Mn}) = pr) \quad (13)$$

This means that agents Ag_{Lc}, Ag_{St} have to control students' progress.

The agents' desires are formulated as follows

$$\langle\langle Ag_{Lc}, Ag_{St}\rangle\rangle\ desire\ A\ \ G\ pr \neq low \wedge \neg danger. \quad (14)$$

That is, the goal of agents is to Ag_{Lc}, Ag_{St} ϵ prevent the student from falling into the risk zone.

For agents $Ag_{Lc}, Ag_{St}, Ag_{Gs}$ you can build the following plans

$$\langle\langle Ag_{Lc}, Ag_{St}, Ag_{Gs}\rangle\rangle\ intend\ A\ \ G\ (pr = medium \vee pr = high). \quad (15)$$

This means that the agents plan to help students gain a sufficiently high level of knowledge.

4. Modeling the agent's function of detecting negative incidents in the educational process

In this paper, the authors emphasized the implementation of the learning progress control function. For this purpose, the agent, in case of detecting negative aspects (student's absence from classes is more than 20%, or the student received the sum of points based on the results of the final/semester academic performance check in accordance with the schedule of the educational process is less than the minimum allowable), sends a message that this student is at risk. An agent is created by adding a corresponding component to the program workspace. Creating an agent includes two actions: programming the agent and linking the created agent to other agents. Linking occurs by connecting two agents in the workspace. Communication between agents can be uni- or bi-directional. In the first case, messages can be transmitted only in one direction and in the other - agents have the ability to exchange messages in both directions. Programming agents and organizing their interaction is done with the help of basic HTML, CSS, JavaScript, PHP, and MySQL capabilities. The monitoring subsystem uses a web camera to track the presence of students in class [28]. Based on the monitoring results, the detection time (in minutes) of the student's presence during the entire class is recorded. This data is stored on the server. According to the set timer parameters (Fig. 4), Student Agent checks the attendance data stored on the server and builds a table (Fig. 6).

unique_id	data	created_at
827ccb0eea8a706c4c34a16891f84e7b	{ "status": "OK", "result": { "827...	2023-02-04
01cfcd4f6b8770febfb40cb906715822	{ "status": "OK", "result": { "...	2023-02-11
dcdcce4af533fdf8581777e1c954a072	{ "status": "OK", "result": { "...	2023-02-18
715a097049577be8cf37147e127de682	{ "status": "OK", "result": { "...	2023-02-25

Fig. 6. Table of student attendance at classes

The data is stored in a table that consists of the following elements: unique_id - a unique record identifier that is generated from the time the record is created and encoded in md5 format; data - information about student attendance, which is stored in json format; create_at - date of creation of the attendance record in the usual date format. The data in the table is updated every week of the training plan. Each record contains information about the attendance of one week.

The results of the analysis of students' success from the course at the time of the inspection are shown in Table 1 (for the period 2023-03-13 – 2023-03-13) and the results of the general analysis of students' success according to the educational plan at the time of the inspection are shown in Table 2, where Th_L – the total hours of lectures; H_L – the hours of the student's attendance at lectures, Th_{lp} – the total hours of laboratory/ practical works; H_{lp} – the hours of the student's attendance at laboratory/ practical works; P_{Min} – the minimum allowable points per semester for the course; S_p

– the student's accumulation of points for the course; N_u – the number of uncompleted laboratory/practical works.

Table 3

The results of the analysis of students' success from the course

Username	Group	Th_{lp} / H_{lp}	P_{Min} / S_p
Albert Joref	KIT-22-1	2/2	20/25
Kirill Petrtenko	KIT-22-1	2/1,8	20/20
Pavlo Sydorenko	KIT-22-1	2/1,5	20/0
Denis Vashenko	KIT-22-1	2/0	20/0

Analysis is performed weekly.

Table 2

The results of the general analysis of students' success according to the educational plan

Username	Group	Th_L / H_L	Th_{lp} / H_{lp}	N_u
Albert Joref	KIT-22-1	50/48	28/26	0
Kirill Petrtenko	KIT-22-1	50/41,8	28/24	1
Pavlo Sydorenko	KIT-22-1	50/36,5	28/20	2
Denis Vashenko	KIT-22-1	50/1,5	28/0	7

Analyzing the diagram in Figure 7 helps to identify the links between different indicators and draw conclusions about student performance and student achievement. For example, the relationship between attendance and academic performance: if students who attend more lectures and labs also have fewer incomplete assignments, we can conclude that regular attendance can affect a student's academic performance. If you click on a row, you can see more detailed information in a separate window (Fig. 8):"status" - is an automatic check of the json file for correctness of formation; "result" - information about the created record: unique key; date of creation; period for which the data was collected; identifier of the study group; information about each student of this group (their attendance at the disciplines that are mandatory in the curriculum). The json (JavaScript Object Notation) format is independent of the implementation language, so it is convenient to use it for data exchange.

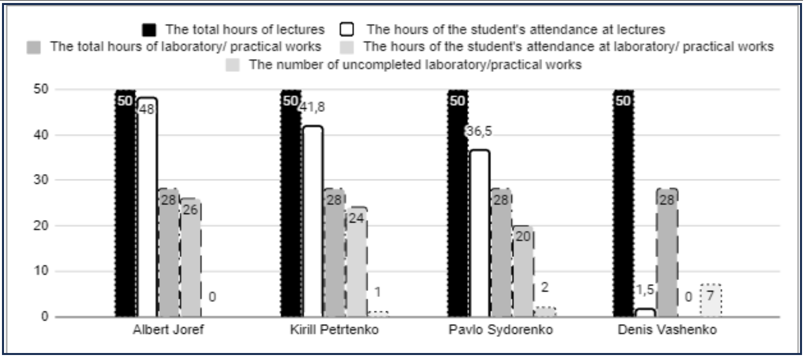


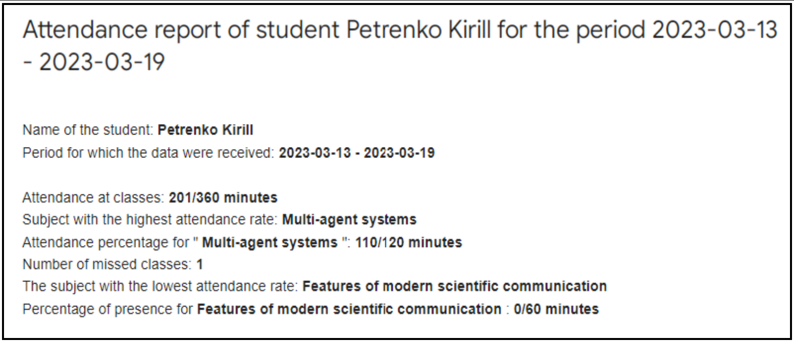
Fig. 7. The diagram of student success for the semester


```
"status": "OK",
"result": {
    "827ccb0eea8a706c4c34a16891f84e7b": {
        "id": "827ccb0eea8a706c4c34a16891f84e7b",
        "date_generate": "2023-03-19 21:00:01",
        "info_from": "2023-03-13 to 2023-03-19",
        "group": "CITm-22-1",

        "students_info": {
            "class_attendance": {
                "Tatarnykov Andrii": {
                    "visits_minutes": "357/360",
                    "subj_best_attendance": "Multi-agent systems",
                    "minutes_best_attendance": "175/180",
                    "missed_classes": "0",
                    "subj_worst_attendance": "Intelligent data analysis",
                    "minutes_worst_attendance": "115/120"
                },
                "Petrenko Kirill": {
                    "visits_minutes": "201/360",
                    "subj_best_attendance": "Multi-agent systems",
                    "minutes_best_attendance": "110/120",
                    "missed_classes": "1",
                    "subj_worst_attendance": "Features of modern scientific communication",
                    "minutes_worst_attendance": "0/60"
                },
                "Sydorenko Pavlo": {
                    "visits_minutes": "140/360",
                    "subj_best_attendance": "Multi-agent systems",
                    "minutes_best_attendance": "90/120",
                    "missed_classes": "3",
                    "subj_worst_attendance": "Intelligent data analysis",
                    "minutes_worst_attendance": "0/120"
```


Fig. 8. Information on student attendance in JSON data exchange format

Depending on the debugging parameters of the developed system, the agent sends a message to e-mail (Fig. 9, 10).

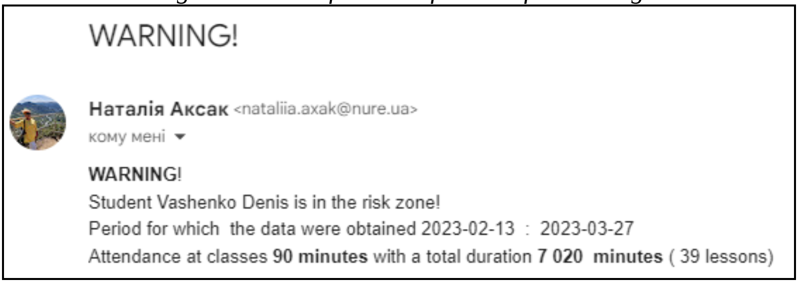


Attendance report of student Petrenko Kirill for the period 2023-03-13 - 2023-03-19

Name of the student: **Petrenko Kirill**
Period for which the data were received: **2023-03-13 - 2023-03-19**

Attendance at classes: **201/360 minutes**
Subject with the highest attendance rate: **Multi-agent systems**
Attendance percentage for " **Multi-agent systems** ": **110/120 minutes**
Number of missed classes: **1**
The subject with the lowest attendance rate: **Features of modern scientific communication**
Percentage of presence for **Features of modern scientific communication** : **0/60 minutes**

Fig. 9. Mail notification of successful training



WARNING!

Наталія Аксак <nataliia.axak@nure.ua>
кому мені

WARNING!
Student Vashenko Denis is in the risk zone!
Period for which the data were obtained 2023-02-13 : 2023-03-27
Attendance at classes **90 minutes** with a total duration **7 020 minutes** (39 lessons)

Fig. 10. Mail notification in case of a negative aspect

Depending on the settings, the notification can be sent weekly in any case (Fig. 8), or the notification will be sent only if negative aspects occur. Figure 9 shows an example when a student Vashenko Denis attended classes for only 90 minutes from the beginning of the semester 2023-02-13 to the current date 2023-03-27, although during this period 39 lessons were held for a total of 7020 minutes. Therefore, this student is at risk of being expelled.

5. Discussion

The approach proposed on the basis of the agent-based learning platform model allows solving the complex problem of timely provision of knowledge together with monitoring of student performance by both teachers and parents. This helps to increase the effectiveness of distance learning. The chosen learning strategy allows us to detect the lack of student progress in a timely manner by monitoring the information available on Moodle and the university's web portal. In the event of negative aspects (prolonged absence of a student from classes, the amount of points received based on the results of the final/semester academic performance check in accordance with the schedule of the educational process is less than the minimum allowable), the system notifies the student, his/her parents and the lecturer that such a student is at risk. Thus, the student has the opportunity to control his or her academic performance and, if necessary, to liquidate his or her debts in a timely manner, i.e., to successfully complete the course and not be expelled. ATL provided a powerful mechanism for expressing the properties of agents, allowed to model agent strategies, i.e. ways of their interaction; express and analyze the temporal aspects of communication and interaction between agents, including the order of events and the time in which they occur.

6. Conclusion

The paper solves an urgent problem that arises during distance learning in an e-learning environment, namely, analyzing and monitoring student performance. To effectively manage the internal learning processes, we propose a model of an agent-based learning platform that can provide services to university management, lecturers, students, and their parents. The novelty of the proposed model lies in taking into account the interconnection of all its components: a multi-agent system, a university web portal, a Moodle learning management system, a cloud data storage, and a subsystem for monitoring student behavior. The multi-agent system is designed to model the process of communication and coordination of participants in the educational process.

The role of each actor of the multiagent system is defined..The use of ATL for the formal specification of a large number of agents performing different roles has allowed us to create a) modal properties of the agent-based learning platform, indicating the main aspects and characteristics of the agents used in the proposed system; b) temporal properties, indicating aspects of time and the temporal nature of agent behavior. The use of the JSON data exchange format made it possible to realize communication between agents. Modeling the identification of negative aspects in the learning process with the help of agents confirmed the feasibility of the proposed approach.

7. References

[1] M. U. Bokhari, S. Ahmad, Multi-agent Based E-learning Systems: A Comparative Study, in: Proceedings of the 2014 International Conference on Information and Communication Technology for Competitive Strategies, 2014, pp. 1-6. DOI:10.1145/2677855.2677875.

[2] S. Mariani, A. Omicini, Special Issue “Advances in Multi-Agent Systems”: Editorial. Appl. Sci. 2023, 13, 3027. DOI:10.3390/app13053027.

[3] del Río, A., J. Serrano, D. Jimenez, L.M. Contreras, F. Alvarez, A Deep Reinforcement Learning Quality Optimization Framework for Multimedia Streaming over 5G Networks. Appl. Sci. 2022, 12, 10343. DOI:10.3390/app122010343.

[4] H. A. Abu-Alsaad, Agent Applications in E-learning Systems and Current Development and Challenges of Adaptive e-Learning Systems, in: 2019 11th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), 2019, pp. 1-6. DOI: 10.1109/ECAI46879.2019.9042015.

[5] I. Vuković, K. Kuk, P. Čisar, M. Banđur, Đ. Banđur, N. Milić, B. Popović, Multi-Agent System Observer: Intelligent Support for Engaged E-Learning. Electronics, 2021, 10(12), 1370. DOI:10.3390/electronics10121370.

[6] F. Elghibarı, R. Elouahbi, F. El Khoukhi, Dynamic Multi Agent System for Revising E-learning Content Materia. The Turkish Online Journal of Distance Education (TOJDE), 2019, 20(1), 131-144. DOI:10.17718/tojde.522434.

[7] N. Viswanathan, S. Meacham, F. F. Adedoyin, Enhancement of Online Education System by Using a Multi-agent Approach. Computers and Education: Artificial Intelligence, 2022, 3: 100057. DOI:10.1016/j.caeai.2022.100057.

[8] M. Amane, K. Aissaoui, M. Berrada, A Multi-agent and Content-Based Course Recommender System for University E-learning Platforms, in: Motahhir, S., Bossoufi, B. (eds) Digital Technologies and Applications. ICDTA 2021. Lecture Notes in Networks and Systems, vol 211. 2021. pp. 663-672. Springer, Cham. DOI:10.1007/978-3-030-73882-2_60.

[9] A. B. Mourão, J. F. M. Netto, SIMROAA Multi-Agent Recommendation System for Recommending Accessible Learning Objects, in: 2019 IEEE Frontiers in Education Conference (FIE), Covington, KY, USA, 2019, pp. 1-9. DOI:10.1109/FIE43999.2019.9028504.

[10] S. Munawar, S. Khalil Toor, M. Aslam, E. Aimeur, PACA-ITS: A Multi-Agent System for Intelligent Virtual Laboratory Courses. Applied Sciences, 2019. DOI:10.3390/app9235084.

[11] Z. Houhamdi, B. Athamena, R. Abuzaineddin, M. Muhairat, A Multi-Agent System for Course Timetable Generation. TEM Journal 8 211 (2019) 211-221.

[12] T. Ivanova, V. Terzieva, K. Todorova, An Agent-Oriented Architecture for Strategy-Based Personalized E-Learning, in: 2021 Big Data, Knowledge and Control Systems Engineering (BdKCSE), Sofia, Bulgaria, 2021, pp. 1-8. DOI: 10.1109/BdKCSE53180.2021.9627260.

[13] M.L. Morales-Rodríguez, J.A. Ramírez-Saldivar, J.P. Sánchez-Solís, A. Hernández-Ramírez, Design of an Intelligent Agent for Personalization of Moodle's Contents. Res. Comput. Sci., 56 (2012) 11-17.

[14] M.M. Al-Tarabily, R.F. Abdel-Kader, G.A. Azeem, M. I. Marie, Optimizing Dynamic Multi-Agent Performance in E-Learning Environment," in: IEEE Access, vol. 6, 2018, pp. 35631-35645. DOI: 10.1109/ACCESS.2018.2847334.

[15] T.O. Almeida, J.F. De Magalhães Netto, A.M.M. Lopes, Multi-Agent System for Recommending Learning Objects in E-Learning Environments, in: 2021 IEEE Frontiers in Education Conference (FIE), Lincoln, NE, USA, 2021, pp. 1-5. DOI: 10.1109/FIE49875.2021.9637258.

[16] A. Alexandru, E. Tirziu, E. Tudora, O. Bica, Enhanced Education by Using Intelligent Agents in Multi-Agent Adaptive e-Learning Systems. Studies in Informatics and Control 24 1 (2015) 13-22.

[17] O.M. Salazar, D.A. Ovalle, N.D. Duque, Adaptive and Personalized Educational Ubiquitous Multi-Agent System Using Context-Awareness Services and Mobile Devices, in: Zaphiris, P., Ioannou, A. (eds) Learning and Collaboration Technologies. LCT 2015. Lecture Notes in Computer Science, vol 9192. Springer, Cham, 2015, pp. 301-312. DOI:10.1007/978-3-319-20609-7_29.
[18] H. El Fazazi, M. Elgarej, M. Qbadou, K. Mansouri, Design of an Adaptive E-Learning System based on Multi-Agent Approach and Reinforcement Learning 11 1 (2021) 6637–6644. DOI:10.48084/etasr.3905.
[19] S. Hammami, H. Mathkour, Adaptive E-learning System Based on Agents and Object Petri Nets (AELS-A/OPN). Computer Applications in Engineering Education 23 2 (2015) 170-190. DOI:10.1002/cae.21587.
[20] N. Tahiri, M. El Alami, An Intelligent E-Learning System for Autistic Children: Multi-Agent Architecture, in: Ezziyyani, M. (eds) Advanced Intelligent Systems for Sustainable Development (AI2SD'2019). AI2SD 2019. Advances in Intelligent Systems and Computing, vol 1102. Springer, Cham. 2020, pp. 83-90. DOI:10.1007/978-3-030-36653-7_8.
[21] F.Z. Lhafra, O. Abdoun, Integration of Evolutionary Algorithm in an Agent-oriented Approach for an Adaptive E-learning. International Journal of Electrical and Computer Engineering 13 2 (2023) 1964-1978. DOI:10.11591/ijece.v13i2.pp1964-1978.
[22] M. Amer, H. Aldesoky, Building an E-Learning Application Using Multi-agents and Fuzzy Rules. Electronic Journal of e-Learning 19 3 (2021) 199-208. DOI:10.34190/ejel.19.3.2308.
[23] N.E. Ghouch, M. Kouissi, E.M. En-naimi, Multi-agent Adaptive Learning System Based on Incremental Hybrid Case-based Reasoning (IHCBR), in: Proceedings of the 4th International Conference on Smart City Applications 2019, pp. 1-6. DOI:10.1145/3368756.3369030.
[24] J.H. Ko, C. Hur, H. Kim, A Personalized Mobile Learning System Using Multi-agent, in: Dean, M., et al. Web Information Systems Engineering – WISE 2005 Workshops. WISE 2005. Lecture Notes in Computer Science, vol 3807. Springer, Berlin, Heidelberg. 2005. pp. 144-151. DOI:10.1007/11581116_15.
[25] D. G. De La Iglesia, J. F. Calderón, D. Weyns, M. Milrad, M. Nussbaum, A Self-Adaptive Multi-Agent System Approach for Collaborative Mobile Learning, in: IEEE Transactions on Learning Technologies 8 2 (2015) 158-172. DOI:10.1109/TLT.2014.2367493.
[26] C.N. Udanor, O.U. Oparaku, A Performance Evaluation of a Multi-Agent Mobile Learning System. International Journal of Interactive Mobile Technologies 10 2 (2016) 10-18. DOI:10.3991/ijim.v10i2.4873.
[27] K. Chouchane, O. Kazar, A New Cloud-Based Adaptive Mobile Learning System Using Multi-Agent Paradigm. International Journal of Organizational and Collective Intelligence (IJOCI) 12 3 (2022) 1-9. DOI:10.4018/IJOCI.306694.
[28] N. Axak, A. Tatarnykov, The Behavior Model of the Computer User, in: IEEE 17th International Conference on Computer Sciences and Information Technologies (CSIT), 2022, pp. 458-461. DOI:10.1109/CSIT56902.2022.10000499.
[29] A. Francalanza, L. Aceto, A. Ingólfsdóttir, Monitorability for the Hennessy–Milner Logic with Recursion. Formal Methods in System Design, 51 1 (2017) 87-116. DOI:10.1007/s10703-017-0273-z.
[30] A. Qasim, S. A.R. Kazmi, I. Fakhir, Executable semantics for the formal specification and verification of E-agents. Indian Journal of Science and Technology, 8 16 (2015) 1-8. DOI: 10.17485/ijst/2015/v8i16/55160.
[31] D. Zanarini, C. Luna, L. Sierra, Alternating-Time Temporal Logic in the Calculus of (Co)Inductive Constructions, in: Gheyi, R., Naumann, D. (eds) Formal Methods: Foundations

and Applications. SBMF 2012. Lecture Notes in Computer Science, vol 7498. Springer, Berlin, Heidelberg. DOI:10.1007/978-3-642-33296-8_16.

УДК 004.89

РОЗРОБКА НАВЧАЛЬНОЇ ПЛАТФОРМИ НА ОСНОВІ АГЕНТІВ З СПЕЦИФІКАЦІЯМИ ТЕМПОРАЛЬНОЇ ЛОГІКИ

Dr.Sci. Аксак Н., Ph.D. Кушнарьов М., Татарников А.

Харківський національний університет радіоелектроніки, Україна
E-mail: nataliia.axak@nure.ua, maksym.kushnarov@nure.ua, andrii.tatarnykov@nure.ua

Анотація. *Розроблено модель агентної навчальної платформи (ALP), в якій на основі поєднання агентної технології та комп'ютерного зору запропоновано підхід, що дозволяє поєднати колективний контроль, координацію та кооперацію агентів для надання освітніх онлайн-послуг. . Запропонований підхід дозволяє, по-перше, учням бути в курсі ходу власної навчальної діяльності, а по-друге, сповістити батьків у разі виникнення негативного інциденту, тобто коли їхня дитина не відвідує заняття за розкладом, встановленим у цьому закладу освіти, або цей учень отримав за результатами підсумкової/семестрової перевірки навчальних досягнень бал згідно з графіком навчального процесу нижче мінімально допустимого. Надання таких послуг реалізується через підсистему моніторингу поведінки студентів та спільноту інтелектуальних агентів. Формальна специфікація агентів була зроблена за допомогою часової логіки змінного часу (ATL). Формалізація використовується для моделювання одночасної поведінки системи з необмеженою кількістю агентів і станів, а також для перевірки деяких властивостей, виражених у вигляді формул ATL. Розроблено програмний прототип агентного середовища для інтеграції розподілених агентів в єдиний інформаційний простір, який поєднує в собі навчальну платформу Moodle, веб-портал університету, підсистему моніторингу поведінки студентів і хмарне сховище Amazon S3.*

Ключові слова*: система управління навчанням, моніторинг, Moodle, навчальний процес, агенти, мультиагентна система, специфікації часової логіки, змінна тимчасова логіка*

Section 9. Control information system

UDC 681.51:004.896

INTERNAL COMBUSTION ENGINE DETONATION ELIMINATION THROUGH THE APPLICATION OF BANDPASS FILTERS IN THE CONTROL INFORMATION SYSTEM

Tykhon Sytnikov

Odessa Polytechnic National University, Ukraine
ORCID: 0000-0002-0137-5643
E-mail: tykhon.sytnikov@gmail.com

Abstract. *The paper is considered approaches to eliminating detonation of an internal combustion engine. A lack of basic solutions has been identified; to improve the efficiency and economy of the system, it is proposed to use a system of equations based on changing the excess air coefficient. This made it possible to reduce hardware costs, overall dimensions, and, as a result, reduce the cost of the system as a whole. The main advantages of this system: versatility, increased engine efficiency and the use of one standard band-pass digital filter. The idea of control when changing the excess air coefficient is as follows. When detonation occurs, the control system increases the excess air coefficient, which affects engine operation in the lean modes of working mixtures. In this case, the engine power decreases, which leads to a decrease in the load on the engine. It leads not only to the toxicity of detonation, but also to fuel economy and a decrease in CO, NOX and CH. The main part of a bandpass filter is to recognize knock and natural engine noise at different frequencies. Due to the large amount of noise in the signal, in addition to its restructuring, it is proposed to use low-order filters of the same type connected in series to increase the steepness characteristics of the filter and narrowing the passbands. The resulting ratio allows calculate frequencies of nth connection of the same type accurately. Bandpass and notch filters are considered. The behavior of the amplitude frequency characteristic when connected in series is described. This solution made it possible to reduce errors in the operation of the detonation detection unit at the initial stages of detonation, which allows for prompt decisions to be made to eliminate unwanted detonation.*
Keywords: *excess air coefficient, knock sensors, signal processing, ignition timing, Industry 4.0-5.0, digital bandpass filter, same type digital bandpass filter, serial connection, amplitude frequency characteristic, phase frequency characteristic, notch filter.*

1. Introduction

The current state of development of hybrid and electric cars does not remove the problems of increasing the efficiency of a gasoline internal combustion engine (ICE) and reducing toxic gases in car exhaust, which is one of the important tasks in the development and operation of a car. The incompatibility of these indicators leads to complication of the engine information and control system. This issue is considered in the work [1].

Increased efficiency depends on the value of the excess air coefficient α

$$\alpha = \frac{G_B}{G_T L},$$

where G_T – weight of fuel supplied to the internal combustion engine cylinders,

G_B – weight of air supplied during the intake stroke to the internal combustion engine cylinders,

L – stoichiometric composition of the fuel (combustible) mixture (14.5 kg of air per 1 kg of fuel).

For engine operation in fuel economy mode, the excess air coefficient α must have the value α = 1.1 - 2.0, and for engine operation in full power mode N_e^{max} and maximum torque M_e^{max} - α = 0.8 - 0 ,9, which increases the engine efficiency (η_e).

However, when the engine operates in the mentioned modes and when low-quality fuel is used, combustion modes with high speeds of the order of magnitude appear in the engine cylinders (1000-2000) m/s, which lead to wear of the working surfaces of the piston group and gas distribution mechanisms, Figure 1. These combustion modes, called detonation, are characterized by the following parameters: detonation speed; mass velocity of reaction products behind the shock wave; distributed temperature, pressure and density of the mixture directly behind the shock wave; thermodynamic properties of fuel and its reaction products, depending on physical and chemical properties.

Measurement of detonation velocity is carried out by photographic cameras, which allow continuous recording of the movement of the flame of gaseous reaction products and the shock waves that occur in the middle and around the detonating charge. This makes it possible to study the processes of chemical reactions that cause detonations. To eliminate these modes, different approaches are proposed. For example, there is a proposal to cool the fuel-air mixture before feeding it into the engine cylinders or to treat this mixture with an acoustic wave [2, 3, 4]. But more often this problem is eliminated by reducing the ignition timing or reducing the fuel supply. So, for example, in Figure 2 shows indicator diagrams of processes in the cylinders of an internal combustion engine (ICE).

Fig. 1. Wear of the working surfaces of the piston group as a result of detonation

2. Detonation and methods for its elimination

Cars use an electronic injection control system – such as BOSH MotronicME7.1.1(which is quite well known and studied), which makes it possible to realize the high power of the W12(6.0 WR12 48v Figure 3.) engine with minimal fuel consumption by coordinating operating modes with operating conditions. Electronic adjustment of the ignition timing is coordinated with signals from installed knock sensors [4]. Such internal combustion engine is interesting as an instance of engine with high complexity and much power in order to have a lot of adjustments in different working modes to reach different goals as power, fuel economy, get the most clean and eco-friendly exhaust. Small reference about the fuel system, ignition system, engine management of this engine model: two linked common rail fuel distributor rails, multi-point electronic sequential indirect fuel injection with twelve intake manifold-sited fuel injectors; centrally positioned NGK longlife spark plugs, mapped direct ignition with 12 individual direct-acting single spark coils; Bosch Motronic ME 7.1.1 electronic engine control unit (ECU), cylinder-selective knock control via four knock sensors, permanent lambda control, water-cooled alternator[5, 6].

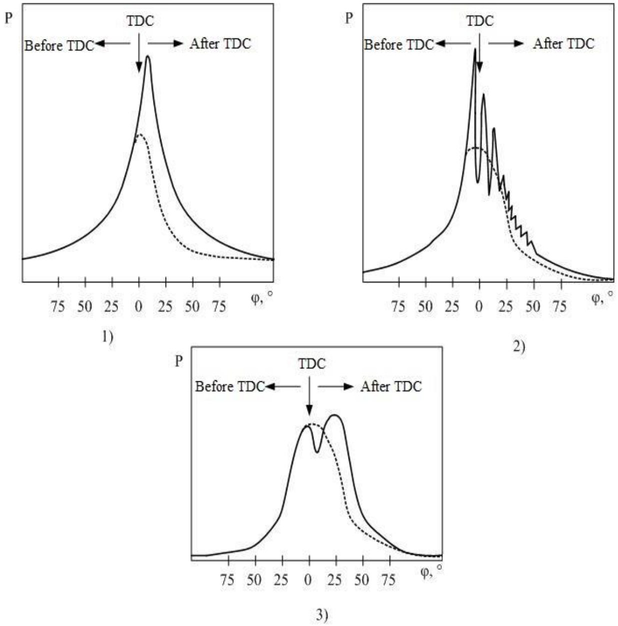


Fig. 2. Indicator diagrams of processes in internal combustion engine cylinders. 1 - with optimal ignition, 2 - during combustion with detonation, 3 - with late ignition, where TDC is top dead center.

This is not the up to date model but this is the predecessor of newer W12 engine models with slight modification, but cost is too high. As an example 6.0 WR12 48v TFSI. Modern regular cars are use BOSCH MEG 17.9.21 which usually used with 1.6 internal combustion engine with significantly cheap car models. In general, a knock sensor is installed on the cylinder block and detects vibrations [1] that occur during detonation. Processing of these signals in the electronic control unit allows, depending on the degree of detonation, commands to change the ignition timing. A general simplified block diagram of knock control that is used in these systems is shown in Figure 4. The signal from the knock sensor at the moment of absence of detonation is sent to one channel of the microcircuit (to detect the engine's own noise), and at the moment of detonation the signal is sent to another channel of the microcircuit.

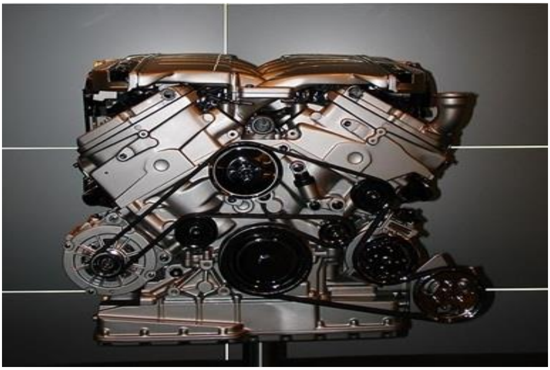


Fig. 3. Engine W12 6.0 WR12 48v

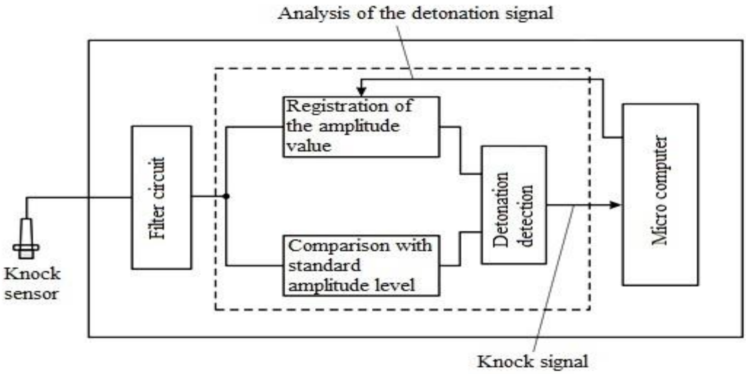


Fig. 4. General simplified block diagram of the knock control system

Both signals are sent to a programmable integrator, in which the detonation signal is released and integrated. The amplified signal enters the ADC, the output of which is supplied to the microcontroller. The microcontroller receives information, processes it and sends control signals to the electronic ignition unit. The microcontroller changes the filter coefficients of the microcircuit in accordance with the engine speed of the crankshaft. The solutions discussed above for eliminating internal combustion engine detonation are based mainly on controlling the ignition timing as the simplest solutions, but not entirely effective and economical.

Elimination of detonation based on changes in the excess air coefficient. Modern development of specialized computer and information control systems makes it possible to build another control system based on changes in the excess air coefficient. In this case, the electronic engine control unit turns into an information and control system (ICS). The development of such a system made it possible to reduce hardware costs, overall dimensions, and, as a result, reduce the cost of the system as a whole. The main advantages of this system: versatility, increased engine efficiency and the use of one single type digital bandpass filter, Figure 5 [4].

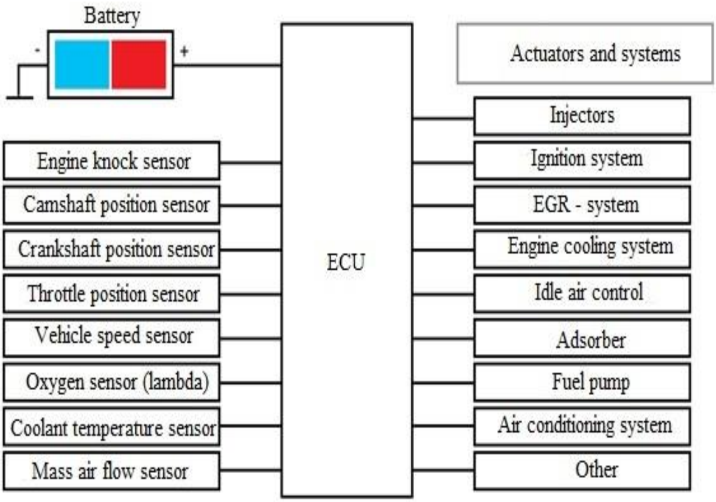


Fig. 5. Engine management information and control system as a further development of the electronic control unit

Detonation is characterized by the appearance of frequencies with high amplitude in the high-frequency part of the spectrum, Figure 2, isolated using a bandpass filter. Detonation detection is performed by subtracting the current signal and the average amplitude of the sensor signal in the absence of detonation.

The degree of detonation is determined by counting the number of amplitudes whose magnitude exceeds the standard value characteristic of the onset of detonation. After detonation is recognized, depending on its degree, the ignition timing is reduced. If after this detonation disappears, the ignition timing is gradually increased, that is, the ignition timing is controlled so that it is close to the detonation edge [1]. The standard engine control units of Impreza cars (Subaru concern) implement an active adjustment system, which allows, in normal engine operation, make corrections to the ignition timing angle of approximately from 3° to +12°. Yes, for new turbocharged control units you can operate with the significantly larger amendments. The use of analog control systems and microcircuits based on them did not improve noise immunity and control quality. To improve these indicators, combined analog-digital control systems are used. For example, to detect detonation in internal combustion engines, automobile manufacturers use knock noise signal processors.

These processors perform analog signal processing, but the use of this microcircuit requires the installation of an external ADC and a microcontroller at its output, which controls the filter coefficients of the microcircuit and produces control signals to the ignition system [4].

The idea of control when changing the excess air coefficient is as follows. When detonation occurs, the control system increases to 1.1÷2.0, this switches the engine operation to a lean mixture mode. In this case, a drop in engine power occurs, which will force the driver to switch to a lower gear, and a car equipped with an automatic transmission will do this itself without reducing the comfort of travel and thereby reduce the load on the engine, which will lead not only to the disappearance of detonation, but also to fuel savings. and reduction of CO, NOX and CH. When constructing such a system, its structural organization is simplified, close to typical, Figure 6 [1]. Based on the commands of the command generation unit (CGU), the multiplexer (MUX) connects the required sensor to the analog to digital convertor (ADC) and the digitized signal enters the information collection unit (ICU). After the information collection unit (ICU), all data collected from the sensors in digital form via the USART protocol is fed to the information processing device. Thus, the device takes data from all sensors used by the common information control system for controlling the coefficient α and the ignition timing.

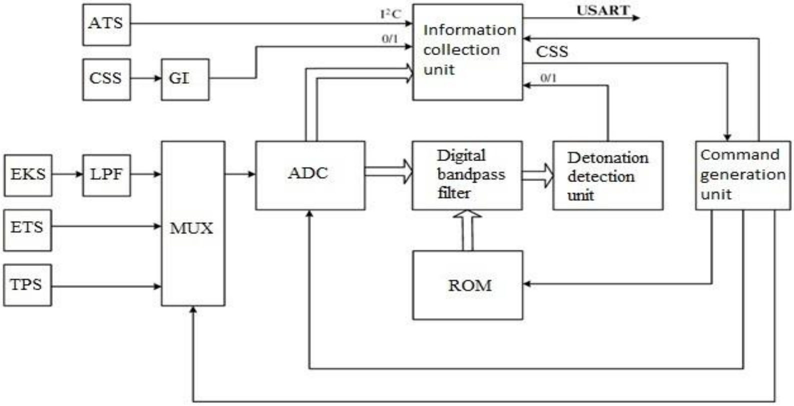


Fig. 6. Generalized block diagram of the detonation tracking system as an integral part of the internal combustion engine information control system. ATS - Air temperature sensor; CSB - crankshaft speed sensor; GIB - galvanic isolation; ICU - information collection unit; EKS - engine knock sensor; ETS - engine temperature sensor; TPS - throttle position sensor; LPF - low pass filter; DBF - digital bandpass filter; DDU - detonation detection unit; CGU - command generation unit; ROM – read only memory; MUX – multiplexor; ADC – analog to digital convertor; USART - universal synchronous and asynchronous receiver-transmitter.

The digital signal from the engine knock sensor (EKS) is fed to the input of the digital bandpass filter (DBF). After filtering the signal, the detonation detection unit (DDU) analyzes the received signal and transmits the result to the information collection unit (ICU). The use of a crankshaft speed sensor (CSS) and engine knock sensor (EKS) made it possible to accurately determine the moments of detonation and the moments of its absence. This made it possible to use a digital filter, which, when there is no detonation, monitors the engine’s own noise, and when detonation occurs, the sum of the detonation signals and the engine's own noise. The difference between these values makes it possible to detect the presence of detonation. It should be noted that the bandpass filter (BPF) must track engine crankshaft rotation speed (CSB). In this case, there is a problem in constructing a tunable bandpass filter. When creating an information and control system for internal combustion engine control, it is necessary to solve a number of issues of processing sensor signals in real time and with limited computing resources. Today, the general approach to building such systems is based on the concept of Industry 4.0. Computerization and informatization of many research processes and industrial production have led to the emergence of the Industrial Internet of Things (IIoT) at cars equipment factories [7]. This direction makes it possible to significantly automate all processes by supplying equipment with multifunctional sensors, actuators and controllers. The collected data is processed in the management information system, which allows you to quickly make informed and informed decisions on adjusting the functional activities of the facility, sometimes automatically.

The role of personnel in this case is reduced to monitoring the operation of systems and responding only to emergency situations to ensure safety and reliability [7-10]. The presence of wireless networks and cloud technologies facilitate the rapid collection of data, which, after primary processing, is sent to the analysis and decision-making center. Further development of such systems is moving towards the humanization of decision-making and friendly contact with people in accordance with the concept of Industry 5.0 [10]. Similar problems are encountered when using various mobile objects, as well as in systems for resolving critical situations [11–15].

Going forward with using IoT in car, general directions of such integration are:

– Direction which is improved car user experience as a passenger such as high-quality in-vehicle infotainment [16].

– Direction which is improve car user experience as a owner\driver, such as over the air update(OTA), effective fleet management, optimal route planning, lower possibility of human mistake, improved driver safety, self driving vehicle.

– Direction which is improve interface between machine – head unit (HU) and driver such as improved vehicle performance and maintenance, personalized driver experience, autonomous driving.

From our point of investigation the last paragraph is most interesting. IoT-equipped vehicles can monitor their own performance and diagnose issues by themselves. Manufacturers receive real-time data about a car’s health. And this allows for proactive maintenance and even remote troubleshooting. For instance, imagine a car’s engine detects a potential problem that can affect your car’s health in the long sun. Now, this IoT-enabled car can send an alert to the owner and the manufacturer. This way, the car’s problem is diagnosed and repaired on time, thus reducing its downtime [16]. Predictive analytics stands out as a remarkable aspect of IoT automotive technology. Within this realm, sensors integrated into various car components gather data and transmit it to a centralized platform. Through the utilization of algorithms, this data undergoes processing, enabling the analysis of future outcomes for each component based on its current performance. Moreover, the IoT automotive maintenance system helps individuals to take proactive measures to prevent sudden breakdowns of their vehicle parts. Similar to the dashboard indicators in a car, this system promptly notifies the driver about potential malfunctions. However, instead of relying solely on in-vehicle alerts, these notifications are transmitted to the driver's mobile device well in advance of any upcoming issue. Consequently, the driver gains the advantage of taking cost-effective and time-saving steps to circumvent component failures while on the road. The predictive maintenance capabilities extend beyond individual vehicles and can be applied to entire fleets as well. This proves particularly beneficial for vehicles that endure lengthy journeys before reaching their destinations, especially those involved in transporting heavy loads. By harnessing the power of the automotive maintenance system, individuals can verify the performance of their vehicles and proactively repair or replace car parts before they succumb to failure. The impact of IoT on the automotive industry has been tremendous. IoT has revolutionized the way vehicles operate, communicate, and interact with the world around them. As the technology continues to evolve and become more widely adopted, we can expect to see even greater changes and advancements in the automotive industry [17]. Technical supervision of the vehicle is an unmet need for users since the speed of the car, the pressure of the tires, and the engine's temperature have direct impacts on the performance of a vehicle. A higher engine's temperature leads to enhanced fuel consumption that increases greenhouse gas emissions. The engine's temperature also affects the efficiency of the turbocharger's compressor and the engine's shut down. IoT can characterize the engine's temperature, speed, and tire pressure to provide quick notifications to the driver. The developed IoT system can collect a variety of data using the most accurate sensor devices to assist the vehicle with fast navigation, crash avoidance, and road condition detection. These features are attractive to users and manufacturers to upgrade the vehicle with up-to-date IoT devices. IoT can develop automotive communications and instantly integrate other devices to assist drivers by collecting traffic data [18]. To create such systems, it is necessary to have filters in the sensor signal processing channel that are capable of restructuring their characteristics by software or hardware, depending on the operating conditions, operating modes and the presence of interference to improve the efficiency of the system as a whole. The digital channel for sensor signals processing in mobile autonomous systems includes various filters. Basically, it is created on low-order filters, because this is due to the low cost of calculating the transfer function coefficients and the number of coefficients, the ease of setting or rebuilding the filter, moderate power consumption and processing time. For example, a first-order Butterworth digital bandpass filter, which is described by a second-order transfer function

$$H(z) = \frac{a_0 + a_1 z^{-1} + a_2 z^{-2}}{1 + b_1 z^{-1} + b_2 z^{-2}}, \quad (1)$$

in the numerator $a_2 = -a_0$, $a_1 = 0$. For the restructuring, it is necessary to calculate one coefficient in the numerator and two in the denominator. There are three coefficients. With the fourth order of the bandpass filter, the transfer function of which is already described by the eighth order

$$H(z) = \frac{a_0 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3} + a_4 z^{-4} + a_5 z^{-5} + a_6 z^{-6} + a_7 z^{-7} + a_8 z^{-8}}{1 + b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3} + b_4 z^{-4} + b_5 z^{-5} + b_6 z^{-6} + b_7 z^{-7} + b_8 z^{-8}}. \quad (2)$$

In this transfer function, the denominator coefficients must be calculated all, and in the numerator the coefficients $a_1 = a_3 = a_5 = a_7 = 0$, and $a_0 = a_8$, $a_6 = -a_2$. For such filter implementation, eight coefficients in the denominator and two coefficients in the numerator must be calculated. It is necessary to calculate 10 coefficients. The amplitude frequency characteristic (AFC) is shown in Figure 7-10.

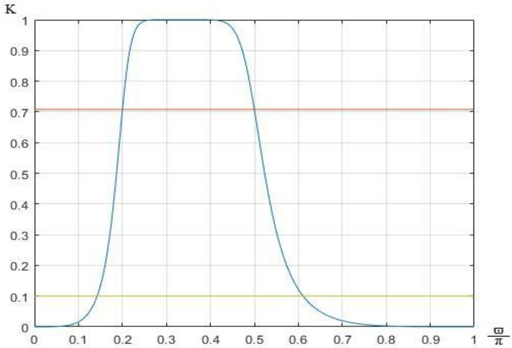


Fig. 7. Amplitude frequency characteristics of digital filters of the fourth order Butterworth

As can be seen from Figure 7-10, fourth-order digital filters have steep edges, however, the presence of ripple in passband and stopband limits the widespread use of such filters. In addition, when working in real time, there are restrictions on the amount and time of calculation, on the restructuring and duration of the transition process. It is known that the higher the order of the filter, the longer the transition process and the problem of filter stability arises, which is already associated with the word length of the representation of the filter coefficients and intermediate calculations. In this regard, it is more convenient to use first-order digital filters, Figure 11. However, it should be noted that all first-order filters have the same amplitude frequency characteristic, but the steepness of the fronts of the amplitude frequency characteristic is low. In the sensor signal processing channel, typical tasks are - changing the cutoff frequencies and bandwidth, as well as the steepness of the fronts of the amplitude frequency characteristic. Therefore, it is advisable to use a first-order filter as a low-order filter. Formulas for tuning its cutoff frequencies and bandwidth are given [19, 20], to increase the steepness of the amplitude frequency characteristic, use a series connection of the same type of filters.

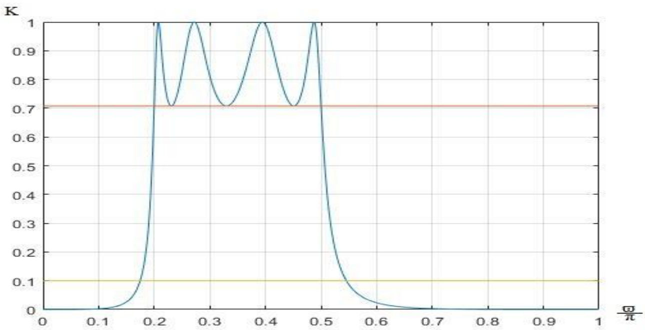


Fig. 8. Amplitude frequency characteristics of digital filters of the Chebyshev of the first

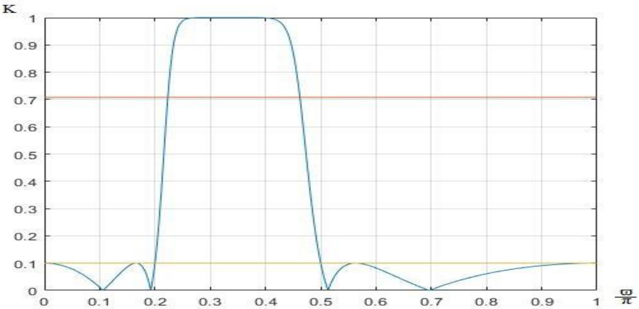


Fig. 9. Amplitude frequency characteristics of digital filters of the Chebyshev of the second kind

3. Series Connections of Same Type Digital Band Pass Filters and Its Effect on Frequency Characteristic

When transfer functions are connected in series, their transfer functions are multiplied

$$H(z) = \prod_{i=1}^{n} H_i(z), \tag{3}$$

where $H_i(z)$, $H(z)$ — i-th transfer functions and the transfer function of the serial connection, respectively.

Since the transfer functions are of the same type $H_o(z)$ and consist of amplitude frequency and phase frequency characteristics, can write

$$H_o(j\overline{\omega}) = H_o(\overline{\omega}) \cdot e^{j\,\phi_o(\overline{\omega})}, \tag{4}$$

where $H_o(\overline{\omega})$, $\phi_o(\overline{\omega})$ — respectively, the amplitude frequency and phase frequency characteristics of the main filter of the same type, $\overline{\omega}$ — where is the normalized angular frequency, $\overline{\omega} = 2\pi \frac{f}{f_d}$, $\overline{\omega} \in [0, \pi]$, f, f_d, — line frequency and sample rate respectively.

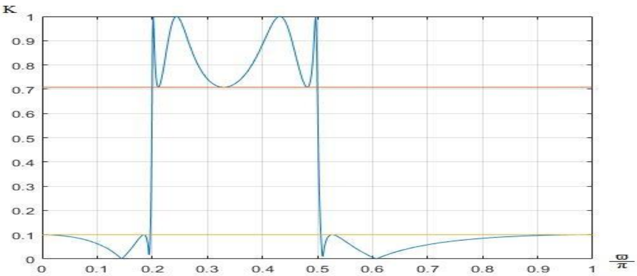


Fig. 10. Amplitude frequency characteristics of digital filters of the Elliptic

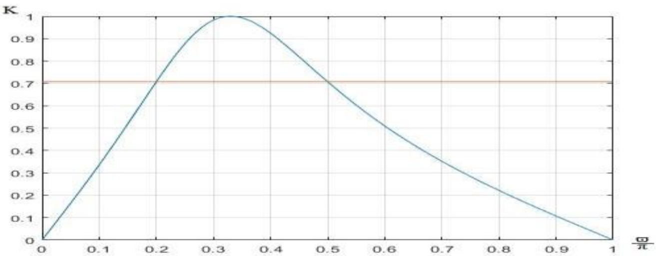


Fig. 11. Amplitude frequency characteristic of a first-order digital Butterworth filter

Then, when similar filters are connected in series, their multiplication is converted into exponentiation

$$H(z) = \prod_{i=1}^{n} H_i(z) = [H_o(z)]^n, \tag{5}$$

and the amplitude frequency characteristic and phase frequency characteristic, respectively, are converted in this way

$$H(j\overline{\omega}) = [H_o(j\overline{\omega})]^n = [H_o(\overline{\omega})]^n \cdot e^{jn\phi_o(\overline{\omega})}. \tag{6}$$

It can be seen from the last ratio that the main changes occur with the amplitude frequency characteristic. It should be noted. that when the components are connected in parallel, their transfer functions add up.

$$H(z) = \sum_{i=1}^{n} H_i(z) = nH_o(z), \tag{7}$$

or for frequency characteristics

$$H(j\bar{\omega}) = nH_o(j\bar{\omega}) = nH_o(\bar{\omega}) \cdot e^{j\phi_o(\bar{\omega})}. \qquad (8)$$

Thus, when connecting the same type of components in parallel, the number of connected components is equivalent to multiplying by the gain factor. In addition, when a component of the same type is included in negative feedback to the same component, the transfer function is obtained

$$H(z) = \frac{H_0(z)}{1+H_0^2(z)}. \qquad (9)$$

With such a connection, the connection order increases, but the shape of the frequency characteristic undergoes significant changes. As a result of the analysis of the connections of the same type of components, it is advisable to use their serial connection in the channel for processing sensor signals in the information control system. Therefore, the aim of the work is to analyze the use of a serial connection of the same type of digital filters and improve the approach for calculating the cutoff frequencies of a new connection. Such a connection makes it possible to increase the efficiency of the information control system in autonomous mobile robotic systems by reducing the bandwidth and increasing the steepness of the amplitude frequency characteristic.

4. Increasing the Steepness of Amplitude Frequency Characteristic of Digital Bandpass Filters

With a series connection of the same type of bandpass filters, the amplitude frequency characteristic of the new connection is compressed, as it were, while the cutoff frequencies are shifted to the center frequency and the steepness of the amplitude frequency characteristic increases, Figure 12.

The transfer function of the main bandpass filter is mathematically described as follows

$$H(z) = \frac{a_0+a_1z^{-1}+a_2z^{-2}}{1+b_1z^{-1}+b_2z^{-2}}, \qquad (10)$$

where a_0, a_1, a_2, b_1, b_2 — the real coefficients of the numerator and denominator, respectively.

When substituting $z^{-1} = e^{-j\varpi}$ or by the Euler formulas $z^{-1} = \cos(\varpi) - j\sin(\varpi)$, ϖ — where is the normalized angular frequency, $\varpi = 2\pi\frac{f}{fd}$, $\varpi \in [0, \pi]$, f, f_d — line frequency and sampling frequency respectively. Based on this transformation, we obtain a complex transfer coefficient, and on its basis, the amplitude frequency characteristic at $a_2 = -a_0$, $a_1 = 0$ and after the transformation we get the square of the amplitude frequency characteristic in the form

$$H^2(\varpi) = \frac{\left(2a_0 sin(\varpi)\right)^2}{(1-b_2)^2+b_1^2+2b_1(1+b_2)cos(\varpi)+4b_2cos^2(\varpi)} \quad (11)$$

It should be noted that the peak frequency of the amplitude frequency characteristic does not change in this case and is determined by the equation, Figure 12

$$\varpi_p = arccos\left(-\frac{b_1}{1+b_2}\right). \quad (12)$$

Usually, the level at which the cutoff frequency is determined is $c = \frac{1}{\sqrt{2}} = 0{,}707$, i.e.

$$H(\varpi_c) = c, \quad (13)$$

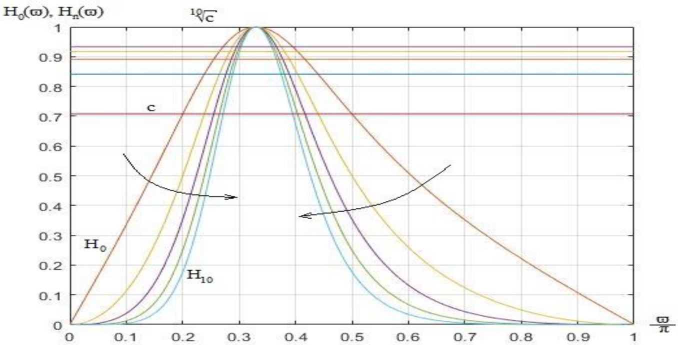


Fig. 12. AFC of digital second-order Butterworth bandpass filters when they are connected in series. Where H_0 - AFC of the first order, H_{10} - AFC of the 10th order, c - the level of the cutoff frequency, $\sqrt[10]{c}$ - the level of the cutoff frequency of the 10th order

Where ϖ_c — amplitude frequency characteristic cutoff at the level c. When multiplying the same type of amplitude frequency characteristic or raising their degree, the level remains the same, but then to determine the cutoff frequencies of the amplitude frequency characteristic of a new connection, when they are connected in series, it is necessary to extract the root of the corresponding order from the level c, i.e $\sqrt[n]{c}$, Figure 13. In

Figure 12, these levels are shown by horizontal lines. In this case, based on the amplitude frequency characteristic of the main filter cutoff frequencies can be calculated for the new compound, Figure 14. On Figure 14 shows the correspondence between the cutoff frequencies of the main amplitude frequency characteristic of the first order at level c and the amplitude frequency characteristic when 5 (five) same-type first-order amplitude frequency characteristic are connected in series.

These cutoff frequencies are determined by the main amplitude frequency characteristic, the parameters of which are known, on a new level. To determine the cutoff frequencies of the new AFC after connecting n filters of the same type according to the main AFC, it is necessary to solve the equation [21]

$$H^2(\varpi) = \frac{\left(2a_0 sin(\varpi_{1n})\right)^2}{(1-b_2)^2+b_1^2+2b_1(1+b_2)cos(\varpi_{1n})+4b_2cos^2(\varpi_{1n})} = \sqrt[n]{c^2} = c^{\frac{2}{n}} \quad (14)$$

where ϖ_{1n} — cutoff frequency at a new level $\sqrt[n]{c}$ according to the main AFC. However, when replacing a_0 with another expression from [19]

$$a_0 = \frac{1-b_2}{2} \quad (15)$$

and instead of $sin(\varpi_{1n})^2 = 1 - cos^2(\varpi_{1n})$, solving this equation, find formulas for determining the cutoff frequencies for the n-th connection of the same type filters. To simplify the representation of the result, introduce new notation

$$A = 4b_2c^{\frac{2}{n}} + (1 - b_2)^2; \quad (16)$$

$$B = -b_1(1 + b_2)c^{\frac{2}{n}}; \quad (17)$$

$$C = (1 - b_2)\sqrt{\left(1 - c^{\frac{2}{n}}\right)\left[(4b_2 - b_1^2)c^{\frac{2}{n}} + (1 - b_2)^2\right]} \quad (18)$$

As a result, obtain the cutoff frequencies of the amplitude frequency characteristic with the n-th connection of the same type of filters, Figure 15.

$$\varpi_{cn1} = arccos(\frac{B+C}{A}); \quad (19)$$

$$\varpi_{cn2} = arccos(\frac{B-C}{A}). \quad (20)$$

In accordance with the formulas obtained, it is possible to determine the bandwidth of such a compound as $\varpi_{BP} = |\varpi_{cn1} - \varpi_{cn2}|$, Figure 16. As can be seen from Figure 16, the bandwidth decreases exponentially. In this case, it is possible to show how many times the bandwidth will decrease with a serial connection, Figure 17. For example, when connecting four components of the same type, the bandwidth is reduced by more than two times, and with eight - more than three times, and ten - the bandwidth is reduced by 3.5 times.

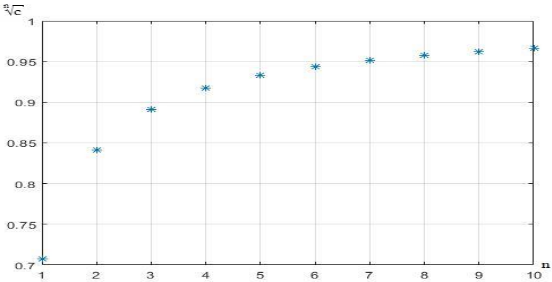


Fig. 13. Graph of the cutoff level dependence on the number of connected filters of the same type

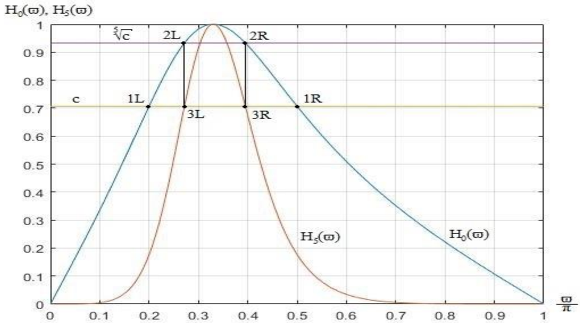


Fig. 14. AFC of the main Butterworth filter of the first order and AFC of the five same-type Butterworth filters of the first order connected into series. Where $H_0(\varpi)$ - AFC of the first order, $H_5(\varpi)$ - AFC of the 5^{th} order, c - the level of the cutoff frequency, $\sqrt[5]{c}$ - the level of the cutoff frequency of the 5^{th} order. Projection 2L and 2R points of the 1^{st} order to 3L and 3R points of the 5^{th} order respectively. 1L and 2R are base cutoff frequency of the first order.

5. Approach Implementation

The implementation of this approach can be both hardware and software. The hardware implementation is based on serial connection of n components of the same type. The main component is calculated based on the data flow and work mode. For example, cutoff frequencies, band pass, gain are determined based on data flow. Block diagram of possible implementation is shown on Figure 18. The scheme proposes a serial connection of several $\mathrm{H_{oi}}$ components of the same type at the same time and registers $\mathrm{Rg_i}$ at the outputs of these components. The MX switch commutes the outputs of the registers to the output of a device. This allows to reduce the time for re-commutation of the components and the time of the transient process, since the necessary data is already in the registers. Such an approach can be implemented by an FPGA. The software implementation was applied for an ultrasonic obstacle sensor data processing in mobile ICS based on the same type of first-order components. For this, a generalized algorithm of the n-th order of the form

$$y_n = a_0^i x_n + i a_0^{i-1} a_1 x_{n-1} + \frac{i!}{2^{i-2}} a_0^{i-2} a_1^2 x_{n-2} + i a_0^{i-3} a_1^3 x_{n-3} + a_1^i x_{n-4} - b_1\, y_{n-1} - \\ - \frac{i!}{2^{i-2}} b_1^2 y_{n-2} - i b_1^3 y_{n-3} - b_1^4 y_{n-4}, \qquad (21)$$

where a_i and b_1 – coefficients of the numerator and denominator of the first order respectively, i – count of the connected first-order elements, $i = 1, 2, 3, 4$.

Such algorithm was implemented, but turned out to be complex for implementation and operations. It requires additional computation, although some components of this algorithm were pre-calculated and stored in memory. The algorithm was developed and a signal processing program was written according to a different principle for the ATMEL MEGA128 microcontroller. This made it possible to reduce the hardware of the system, since all sensors are connected to the ADC, which is located in the microcontroller, and also to reduce the processing time of data from the ADC because it is on the same data bus with the processor. The signal graph of a first-order bandpass filter is shown in Figure 19.

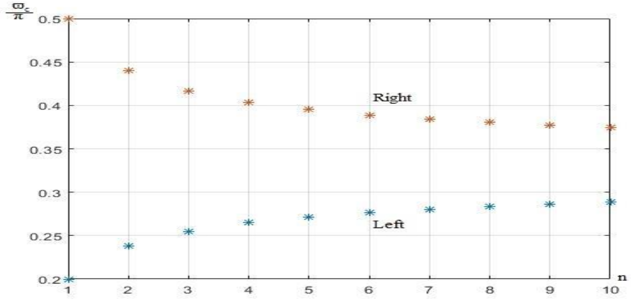


Fig. 15. Graph of the dependence of the cutoff frequencies when the same type filters are connected in series on the number of connections

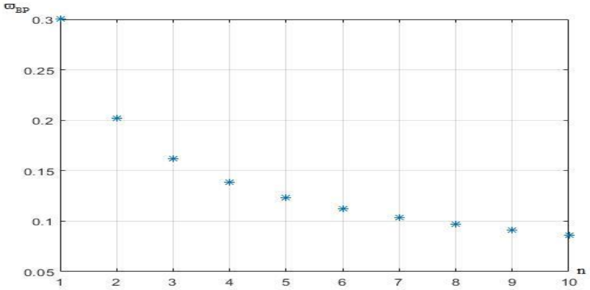


Fig. 16. Plot of bandwidth ϖ_{BP} versus connection number

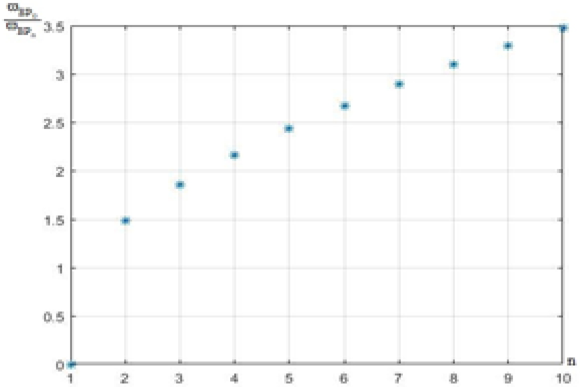


Fig. 17. Bandwidth reduction ϖ_{BP} versus number of connections

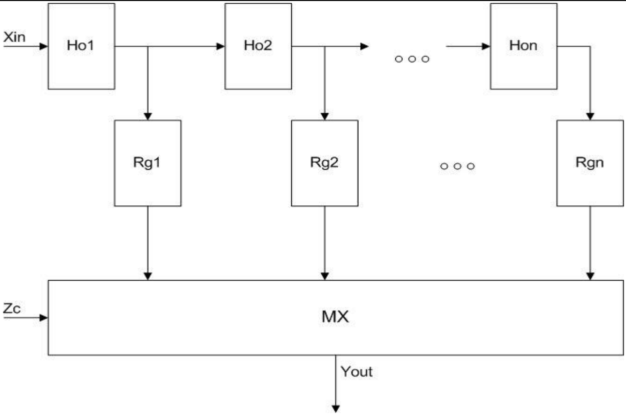


Fig. 18. Block diagram of the serial connection of the same type elements

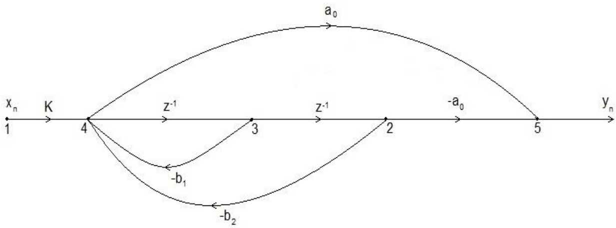


Figure 19: *Signal graph of the first-order digital bandpass filter*

Based on the signal graph a system of equations of the node states ordered signal graph, according which the calculation algorithm was compiled:

$$\begin{cases} x_1[i] = x_n(i); \\ x_2[i] = x_3[i-1]; \\ x_3[i] = x_4[i-1]; \\ x_4[i] = kx_i[i] - b_1x_3[i] - b_2x_2[i]; \\ x_5[i] = a_0 * \{x_4[i] + x_2[i]\} + 2x_3[i]; \\ y_n(i) = x_5[i]; \end{cases} \quad (22)$$

where $x_n, x_i[i], y_n$ – respectively input sequence, state of the graph node, output sequence.

This solution is more convenient for implementation and operation and allowed to reduce the computation time, since some constants were calculated in advance and stored in memory cells.In addition, subroutines were written that were connected if it was necessary to increase the order of the filter and the steepness of the AFC. Modeling fragment of the filter is shown at the Figure 20, 21.

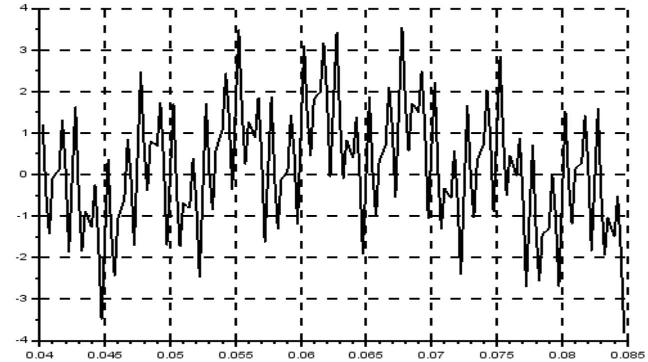


Fig. 20. Fragment of the input signal of the digital bandpass filter

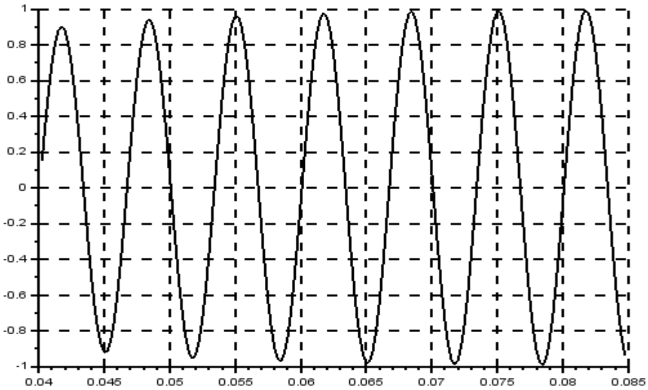


Fig. 21. Fragment of the output signal of the digital bandpass filter

6. Application of This Approach to Notch Filters

Band-pass filters include notch filters, which do not pass the necessary frequencies, but cut them out. By analogy with band-pass filters, in this case there will be compression of the amplitude frequency characteristic. But when connected in series, the amplitude frequency characteristic of notch filters does not narrow, but expands in accordance with the transfer function of this filter.

$$H(z) = \frac{a_0+a_1z^{-1}+a_0z^{-2}}{1+b_1z^{-1}+b_2z^{-2}}. \qquad (23)$$

The mathematical description of the amplitude frequency characteristic of the notch filter looks as this.

$$H(\overline{\omega}) = \sqrt{\frac{(a_1+2a_0sin(\overline{\omega}))^2}{(1-b_2)^2+b_1{}^2+2b_1(1+b_2)\,cos(\overline{\omega})+4b_2cos(\overline{\omega})^2}}. \qquad (24)$$

With the expansion of the amplitude frequency characteristic, its steepness also increases with the expansion of the frequency band cut out from the signal spectrum, Figure 22. This is not convenient in some issues, since with an increase the order it is necessary to narrow the cut-off band with an increase in the steepness of the amplitude frequency characteristic.

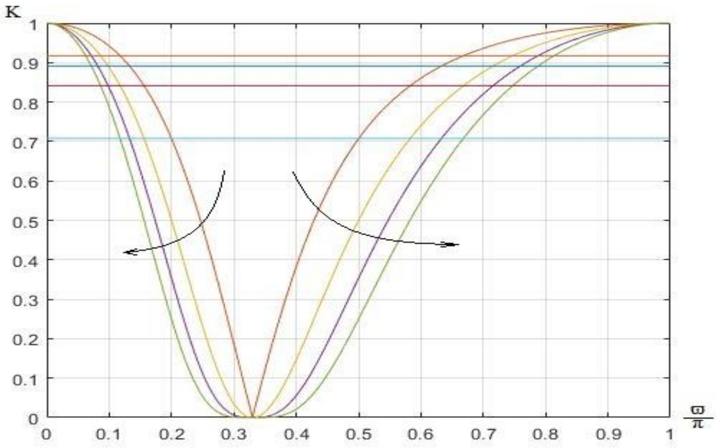


Fig. 22. Amplitude frequency characteristic of first-order Butterworth digital notch filters when they are connected in series

7. Conclusion

Serial connection of frequency-dependent components of low order of the same type leads to exponentiation of the transfer function and AFC. This leads to compression of the AFC with an increase in the steepness of the AFC. The paper shows, using the example of a band-pass filter, that with such a connection, the center frequency does not change. It remains in its place, the cutoff frequencies (left and right) are shifted to the center. In addition, a new approach is obtained for calculating the exact values of the cutoff frequencies and bandwidth.

The advantage of such connection is simple increase in the front of the amplitude frequency characteristic and decrease in the bandwidth of the connection by three times. In the example shown in Figure 8, the bandwidth reduction is - 3.5 times with ten connected filters of the same type. The disadvantages of such connection include not so fast rise and fall of the AFC.

The possibility of implementing this approach on a mobile ICS is shown. This allows on board such systems to calculate the necessary "compression" of the amplitude frequency characteristic, and with limited computing capabilities on board, use a preliminary calculation in the form of a table of values that are stored in memory.

This approach allows you to automatically increase the operational security of data processing in the ICS in the presence of interference.

In addition, the analysis showed that a similar connection of notch filters does not allow obtaining a similar result for processing sensor signals in order to cut out unwanted frequencies in a narrow frequency band. As shown, this results not in a narrowing of the cutting band, but in widening the band, which is not desirable in such systems.

8. References

[1] G. Ukhina, A. Teplechuk, Y. Kiryazov, V. Sitnikov, Control of the band frequency-dependent component by the information and control system to eliminate detonation of the internal combustion engine, Information management systems and technologies. Problems and solutions, Odessa, 2019.

[2] S. Xue, X. Zhang, J. Yang, D. Liu, H. Ma, Changsheng Zhou, Experimental study on the rotating detonation engine based on a gas mixture (2023). DOI:10.3389/fenrg.2023.1136156

[3] Q. Xie, Z. Ji, H. Wen, Z. Ren, Review on the Rotating Detonation Engine and It’s Typical Problems (2020). DOI:10.2478/tar-2020-0024

[4] I. J. Shaw, A.C. Jordan, J. Michael, C. Alfonso, A.M. Ciaran, N.H. Shekh, R. Rubaiyat, C. Rey, R. Paul, A Theoretical Review of Rotating Detonation Engines (2018). DOI: 10.5772/intechopen.90470

[5] Volkswagen Group W-12 engine. [Online]: https://en.wikipedia.org/wiki/Volkswagen_Group_W-12_engine

[6] List of Volkswagen Group petrol engines. [Online]: https://en.wikipedia.org/wiki/List_of_Volkswagen_Group_petrol_engines#6.0_WR12_48v

[7] I. Grangel-González; P. Baptista; L. Halilaj; S. Lohmann; Maria-Esther Vidal; C. Mader, The industry 4.0 standards landscape from a semantic integration perspective, IEEE 22 (ETFA) (2017). DOI:10.1109/ETFA.2017.8247584 2017

[8] B. Santos, Industry 4.0: an overview. 2018. [Online]: https://www.researchgate.net/publication/326352993_Industry_40_an_overview

[9] Industry 4.0. [Online] https://www.cognex.com/ru-ru/what-is/industry-4-0-machine-vision/development.

[10] A. Amr, Future of industry 5.0 in society: human-centric solutions, challenges and prospective research areas, Journal of Cloud Computing 11 (2022). DOI:10.1186/s13677-022-00314-5

[11] S. Semenov, D. Voloshyn, A. N. Ahmed, Mathematical model of the implementation process of flight task of unmanned aerial vehicle in the conditions of external impact, International Journal of Advanced Trends in Computer Science and Engineering, 8 1 (2019). DOI:10.30534/ijatcse/2019/0281.22019

[12] I. Zhuravska, M. Musyienko, O. Tohoiev, Development the heat leak detection method for hidden thermal objects by means the information- measuring computer system. CEUR WorkshopProceedings, (2019) 350-364.

[13] K. Obukhova, T. Savchuk, I. Zhuravska, A. Boiko, V. Nikolskyi, S. Puzyrov, Prevention of unmanned vessels collisions due to pre-modeling the remote control center CPU load, CSIT Proc. Vol. 2. (2021). DOI:10.1109/CSIT52700.2021.9648800

[14] U. Jahn, C. Wolff, P. Schulz, Concepts of a modular system architecture for distributed robotic systems, Journal of Computers, (2019). DOI:10.3390/computers8010025

[15] K. Obukhova, I. Zhuravska, V. Burenko, Diagnostics of power consumption of a mobile device multi-core processor with detail of each core utilization, Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET) (2020). DOI: 10.1109/TCSET49122.2020.235456

[16] V. Amirtham, Connected Vehicles: 10 Ways How IoT Change the Automotive Industry, 2023. [Online]: https://www.contus.com/blog/iot-in-automotive-industry/

[17] Automotive IoT: What is the Impact of IoT in Vehicles?, 2023. [Online] https://www.sotatek.com/automotive-iot-what-is-the-impact-of-iot-in-vehicles/

[18] H. Pourrahmani, A. Yavarinasab, R. Zahedi, A. Gharehghani, M. Hadi, Parisa Bastani, Jan Van herle, The applications of Internet of Things in the automotive industry: A review of the batteries, fuel cells, and engines, (2022). DOI:10.1016/j.iot.2022.100579

[19] H. Ukhina, V. Sytnikov, O. Streltsov, P. Stupen, D. Yakovlev, Transfer function coefficients influence on the processing path bandpass frequency-dependent components' amplitude frequency characteristics properties at the NPP TP ACS, in: 10th International Conference on Dependable Systems, Services and Technologies, DESSERT (2019). DOI: 10.1109/DESSERT.2019.8770050

[20] L.I. Zhyvtsova, H.V. Ukhina, V.S. Sytnikov, O.V. Streltsov, P.V. Stupen, Integrated improvement of the efficiency of computer control system of spatial orientation settings of drilling facilities, Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu (2022). DOI:10.33271/nvngu/2022-1/116.

[21]G. Korn, T. Korn, Mathematical handbook for scietists and engineers, McGraw-Hill Book Company, 1969. DOI:10.1002/zamm.19690490921

УДК 681.51:004.896

УСУНЕННЯ ДЕТОНАЦІЇ ДВИГУНА ВНУТРІШНЬОГО ЗГОРЯННЯ ШЛЯХОМ ЗАСТОСУВАННЯ СМУГОВИХ ФІЛЬТРІВ В ІНФОРМАЦІЙНІЙ СИСТЕМІ КЕРУВАННЯ

Ситников Т.
Національний університет «Одеська політехніка», Україна
E-mail: tykhon.sytnikov@gmail.com

Анотація. *У роботі розглянуто підходи до усунення детонації двигуна внутрішнього згоряння. Виявлено брак основних рішень; для підвищення ефективності та економічності системи запропоновано використовувати систему рівнянь, засновану на зміні коефіцієнта надлишку повітря. Це дозволило знизити витрати на обладнання, габаритні розміри і, як наслідок, знизити вартість системи в цілому. Основні переваги цієї системи: універсальність, підвищена ефективність двигуна та використання одного стандартного смугового цифрового фільтра. Ідея управління при зміні коефіцієнта надлишку повітря полягає в наступному. При виникненні детонації система управління збільшує коефіцієнт надлишку повітря, що позначається на роботі двигуна в збіднених режимах робочих сумішей. При цьому знижується потужність двигуна, що призводить до зниження навантаження на двигун. Це призводить не тільки до токсичності детонації, а й до економії палива та зниження CO, NOX і CH. Основною частиною смугового фільтра є розпізнавання детонації та природного шуму двигуна на різних частотах. У зв'язку з великою кількістю шумів в сигналі, крім його перебудови, пропонується використовувати однотипні фільтри нижчого порядку, з'єднані послідовно, для підвищення крутизни характеристик фільтра і звуження смуг пропускання. Отримане співвідношення дозволяє точно розрахувати частоти n-го з'єднання одного типу. Розглядаються смугові та режекторні фільтри. Описано поведінку амплітудно-частотної характеристики при послідовному з'єднанні. Таке рішення дозволило зменшити помилки в роботі блоку виявлення детонації на початкових етапах детонації, що дозволяє оперативно приймати рішення щодо усунення небажаної детонації.*

Ключові слова: *коефіцієнт надлишку повітря, датчики детонації, обробка сигналів, випередження запалювання, Industry 4.0-5.0, цифровий смуговий фільтр, однотипний цифровий смуговий фільтр, послідовне підключення, амплітудна частотна характеристика, фазочастотна характеристика, режекторний фільтр.*

Section 10. Information systems

UDC 681.391

THE SYNTHESIS OF FAST DCT-II ALGORITHMS USING THE CYCLIC SUBSTITUTIONS TECHNIQUE

Dr. Sci., Prots'ko I. , Rykmas R.

Lviv Polytechnic National University, Ukraine
ORCID: 0000-0002-3514-9265, 0000-0002-1118-2036
E-mail: ihor.o.protsko@lpnu.ua, roman.rikmas@gmail.com

Abstract. *The approach of synthesis the discrete cosine transforms based on cyclic convolutions requires the formation of the block-cyclic structure of the base matrix, an analysis of the block-cyclic structure of the core of the transform and the determination of the minimum number of cyclic convolutions. A model of a compressed description of the block-cyclic structure of the core of a discrete cosine transform in the form of a cyclic decomposition of the substitution is considered. The cyclic decomposition of the substitution contains integer elements of the arguments of the basic function of the transform. In addition, the models are supplemented by a cyclic decomposition of the substitution with simplified values of the elements of the arguments of the basic function and a cyclic decomposition of substitution of signs. An analysis and determination of identical cyclic submatrices in the basic matrix of transform is performed with a variable step of the search, based on the parameters of the model of the discrete cosine transform. As a result of software implementation using the cyclic substitutions technique allows us to speed up the process of analysis of the block-cyclic structure of the core of the transform, reduce the number of cyclic convolutions for computation of the discrete cosine transforms of arbitrary sizes.*
Keywords: *discrete cosine transforms; synthesis of algorithm; analysis block-cyclic structure; cyclic decomposition of substitution*

1. Introduction

The use of a set of effective algorithms for information data processing provides compression, encoding and encryption of input data streams, increases the speed of formation and reliability of transmission of information and communication systems of compact, secure information packets. Thanks to fast algorithms of discrete Fourier class transforms, efficient storage, transmission and processing of multimedia data is achieved with a significant reduction of computational costs [1]. For wide practical application by international standardization organizations, ISO / IEC and ITU-T, 8 types of discrete cosine transforms (DCT I-VIII) are recommended [2]. DCT refers to orthogonal trigonometric transforms that correspond to the properties of discrete Fourier transform (DFT) [3]. Orthogonal trigonometric basic systems of DCT I-VIII provide computation of direct and inverse transforms in the real domain, which is especially important for the effective solution of specific practical problems of audio and video data processing. There are many algorithmic approaches, which can simplify digital transforms and reduce computational cost. For efficient computation of DCT, the various forms of recording the core of transform are used, including matrix multiplication with partial factorization [4], full factorization [5], recursive factorization [6] and other forms [7]. Multivariate effective computing algorithms of Fourier class are divided into: algorithms for size of radix two, split radix, mixed radix, odd size, and prime factors composite transform size. In many scientific fields and several problems, not only fast algorithms, block circulant matrices have been used [8]. Mathematical rules, linear algebra and graph theory have some techniques by which calculation cost and size of the structural matrices can be reduced in repetitive, regular and circulant structures. Symmetry and regularity of structures or repetitive structures have been widely studied in details in book [9]. One of the approaches to the development of efficient algorithms is the ability to compute transforms of Fourier class through cyclic convolutions, which was first described for discrete Fourier transform in the publication by Charles Raider in 1968 [10]. This approach uses fast algorithms for convolutions' computing. The efficient cyclic convolutions will lead to efficient computations of Fourier class transforms. The most well-known ones are the Vinograd algorithms of small sizes, the Agarwal–Cooley algorithm, the algorithm of polynomial transformations [11], the block pseudocirculant algorithm [12], the parisection algorithm (Pitassi) [13], or hardware systolic architectures of convolvers using distributed arithmetic [14]. As a result of the use of flexible arrangement of transforms based on the formed block-cyclic structure of the Fourier transform basis, computational complexity decreases, parallel computing of convolutions is performed, and technical parameters (area, delay) of the computer during the ultra-large integrated circuit designing [15] are improved. Discrete transforms and convolutions are the main operations and key tools in signal processing of system and service engineering. Between of variety the algorithms, leading to efficient computations, have the historical development of fast algorithms of Fourier class using cyclic convolutions. The development the famous approaches of efficient computation of DFT based on cyclic convolutions includes transform:

– for prime size N, where the DFT is trivially are computed the remaining ($N-1$) components using an efficient cyclic convolution algorithm presented by Rader (1968) [10];

– for arbitrary size, where individual discrete components DFT are computed using a cyclic convolution algorithm presented by Goertzel (1968) [16];

– for arbitrary size, where chirp-algorithm define DFT via cyclic-convolution and additional multiplications presented by Bluestein (1970) [17];

– for the integer power of prime size, presented by Winograd Fourier Transform Algorithm (1978) [18].

These algorithms present in the various forms and based the properly rearranging the basic matrix are leaded to pseudo/quasi cyclic structures. The techniques can reform the size of transform to different set of the less sizes the cyclic structures. The paper [19] shows that when the length of a p prime is such that $(p-1)/2$ is odd, the DCT can be computed as two cyclic convolutions, each of length $(p-1)/2$. The paper [20] proposes to decompose the computation of the N point DCT into two matrix-vector multiplications, where each matrix is of size $(M-1)\times(M-1)$ and $M = N/2$. Each of the decomposed matrix-vector products is then converted into a pair of $[(M-1)/2]$ point circular convolution-like operations for reduced-complexity of concurrent systolization. Synthesis of fast algorithms often use a direct method based on the properties of the base matrix of transform and indirectly method, which involved other fast algorithms of discrete transform [21]. Modern computation of discrete transform of Fourier class the signals need to use the generalized scheme of efficient algorithms. For many well-known fast algorithms are applied purely algebraic methods. The mathematical principles establish by the each algorithm and justify its structure. The objective of the paper is study the synthesis of fast DCT-II algorithms using the cyclic substitutions technique. The cyclic decomposition of the substitution is used to bring the harmonic basis of DCT-II to a set of cyclic submatrices. Execution of the synthesis of fast DCT-II algorithms requires an analysis and research of the obtained block-cyclic structure of the core of the transform in order to reduce the computational complexity and efficient organization of its execution. It is a further development of approach for the efficient computation of Fourier class is to bring a harmonic basis to block-cyclic matrix structures and computations of transforms using fast cyclic convolutions [22].

2. DCT of type II as the periodic symmetry extension of DFT

The DCT reflect the input data to linear combination of weighted real basis functions. The DCT are a further improvement of the DFT for real input data, because the complex Fourier transform is redundant in this case. The cosine and sine transforms and Fourier transforms are interconnected by appropriate mathematical relations, which allow us to choose an effective way to compute one transform through another. According to the DFT theory, for trigonometric transforms on the basis of interval [0, π] there is a need to continue input sequence twice. The Fourier-related transforms that operate on a function over a finite domain, such as the DFT or DCT or a Fourier series, can be thought of as implicitly defining an extension of that function outside the domain (Figure 1). That is, once you write a function *f*(*x*) as a sum of sinusoids, you can evaluate that sum at any *x*, even for *x* where the original *f*(*x*) was not specified. The DFT, like the Fourier series, implies a periodic extension of the original function.

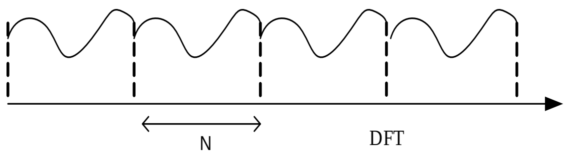


Fig. 1. The periodic extension of the original function for DFT

The axis of symmetry of discrete samples can be on a sample (even number) or between two samples, which corresponds to a shift in the half interval (odd number) sampling. This allows different options of transform under the boundary conditions of real input data [23]. The DCT, like a cosine transform, implies an even extension of the original function (Figure 2). It is equivalent to a DFT of roughly twice the length, operating on real data with even symmetry (since the Fourier transform of a real and even function is real and even). However, the DCTs operate on finite discrete sequences, two issues arise that do not apply for the continuous cosine transform. First, there is a need to specify at both the left and right boundaries of the domain. These different boundary conditions strongly affect the applications of the transform and lead to uniquely useful properties for the various DCT types.

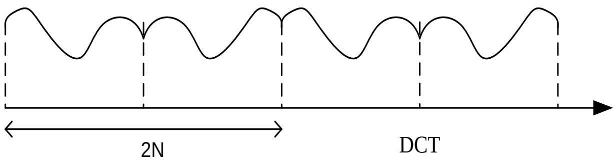


Fig. 2. The even extension of the original function for DCT-II

The most common variant of discrete cosine transform is the type-II DCT, which is often called simply "the DCT". The DCT-II implies the boundary conditions in the case of half interval sampling symmetrically at both the left and right boundaries. The illustrations of the implicit even extensions of the DCT is presented in Figure 3.

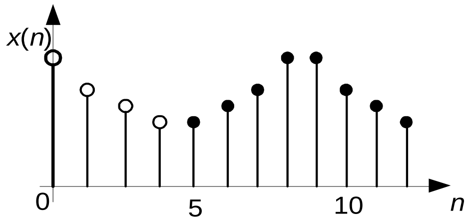


Fig. 3 The illustration of the implicit even extensions of DCT-II for input data N=4

The DCT-II proof is performed by calculating the DFT for the case of periodic continuation of the finite input sequence, which is expressed as a linear combination of complex exponents. While the DFT contains implicit periodicity, the DCT includes both periodicity and parity. For the DCT-II, the sequence $x[n]$ continues to $2N$-periodic as a rule:

$$\tilde{x}_1(n) = x[((n))_{2N}] + x[((-n-1))_{2N}]. \quad (1)$$

Since the overlaps of the extreme samples do not occur, the equality at $n = 0, 1, 2,...,N-1$ is satisfied for the unmodified sequence. This continuation is called the periodic symmetry of the second type (Figure 4), the centres of pair symmetry of the sequence are placed in HS (half samples): -1/2, (N-1/2), (2N-1/2) and etc. That is, the axis of symmetry passes between extreme samples, what corresponds to the set of symmetries given in Figure 4.

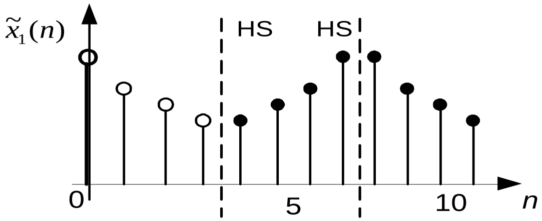


Fig. 4. The continuation of the 4-element sequence corresponding to DCT-II

Let's perform the forward and reverse DCT-II for the extended to 2N-periodic sequence:

$$\tilde{x}(\mathrm{n}) \cong \begin{cases} x(m), \ (0 \le m \le N-1) \\ x(-m-1), \ (-N \le m \le -1) \end{cases} . \quad (2)$$

According to Figure 5 with N=4 element sequence we have a 2N sequence within $-N < n < N$-1, which is periodic with period 2N and is even with respect to the axis m = -1/2

$$x'(m) = x'(-m-1) = x'(2N-m-1) \quad . \quad (3)$$

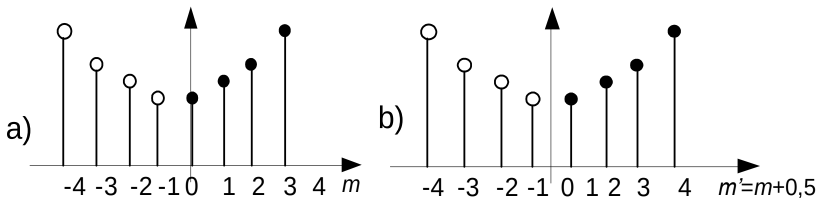


Fig.5. The continuation of the sequence (a) to 2N-periodic and (b) offset by m'= m + ½

If you shift the sequence to the right by ½ according to the Figure 5, b) the sequence index will be $m' = m + 1/2$, then the index $m = m'-1/2$, and accordingly

$$x'(m) = x'(m'-1/2) . \quad (4)$$

Then the representation of the DFT sequence with 2*N* samples for *n*=0,1,..., 2*N*-1 will be

$$\frac{1}{\sqrt{2N}}\sum_{m'=-N+1/2}^{N-1/2} x(m'-1/2)\exp(-j2\pi m'n/2N)], \tag{5}$$

$$\frac{1}{\sqrt{2N}}\sum_{m'=-N+1/2}^{N-1/2} x(m'-1/2)\cos(2\pi m'n/2N) - \\ -\frac{j}{\sqrt{2N}}\sum_{m'=-N+1/2}^{N-1/2} x(m'-1/2)\sin(2\pi m'n/2N). \tag{6}$$

In the obtained expression we have $x'(m) = x'(m'-1/2)$ that is even, $\cos(2\pi m'n/2N)$ and $\sin(2\pi m'n/2N)$ are respectively even and odd functions up to $m'= 0$ or $m'= -1/2$. Therefore, the first sum is doubled within $m'= 1/2,\ldots,N-1/2$, and the second sum will be equal to zero:

$$\frac{1}{\sqrt{2N}}\sum_{m'=-N+1/2}^{N-1/2} x(m'-1/2)\cos(2\pi m'n/2N) = \\ = \sqrt{\frac{2}{N}}\sum_{m'=1/2}^{N-1/2} x(m'-1/2)\cos(2\pi m'n/2N) \tag{7}$$

Substituting $m'=m+1/2$, we obtain DCT-II for $n=0, 1, 2,\ldots,N-1$:

$$X^{c2}(n) = \sqrt{\frac{2}{N}}\sum_{m'=1/2}^{N-1/2} x(m'-1/2)\cos(2\pi m'n/2N) = \\ = \sqrt{\frac{2}{N}}\sum_{m=0}^{N-1} x(m)\cos[(2\pi(2m+1)n/2N] = \sum_{m=0}^{N-1} c(n,m)x(m), \tag{8}$$

where $c(n,m)$ is the basis function of DCT-II.

The output value of $X^{c2}(n) = X^{c2}(-n)$ is even with the period 2*N*

$$X^{c2}(N+n) = X^{c2}(N+n-2N) = X^{c2}(n-N) = X^{c2}(N-n). \tag{9}$$

Therefore, $n = 0,1,2,\ldots,2N-1$ the number of output values is reduced by half $n = 0,1,2,\ldots, N-1$.

Direct DCT-II can be represented by the formula:

$$X^{c2}(n) = a(n)\sum_{m=0}^{N-1} x(m)\cos[\frac{n(2m+1)\pi}{2N}] = \sum_{m=0}^{N-1} c(n,m)x(m), \quad n = 0,1,...,N-1. \tag{10}$$

All rows $c(n,m)$ of vectors $[c(k,0),\ldots,c(k,N\text{-}1)]$ are orthogonal and normalized except the first:

$$\sqrt{\sum_{m=0}^{N-1} c^2(n,m)} = \sqrt{\frac{2}{N}\sum_{m=0}^{N-1} cos^2[(2m+1)n\pi/2N]} = \begin{cases}\sqrt{2}, & n=0\\ 1, & n=1,2,...,N-1\end{cases} . \tag{11}$$

The coefficient makes them orthonormal

$$a(n) = \begin{cases}\sqrt{1/N}, & n=0\\ \sqrt{2/N}, & n=1,2,...,N-1\end{cases} .$$

DCT-II is widely applied for several reasons. Firstly, basis of DCT functions are well approximated to Karhunen-Loyeva transfer function for a big number of stationary stochastic processes, which allows presenting the signal of a given accuracy with a minimal number of components. Secondly, DCT-II is included as part of some efficient DFT algorithms such algorithm. Thirdly, DCT contains a number of special properties, because the conversion is concentrated in the lower indices and more intense and zeroing the remaining output values does not lead to a significant loss of signal energy that prevents edge effects at the block encoded images [24]. DCT-II is used in many applications, especially in processing digital signals audio and video.

3. Mathematical model of the of the block-cyclic structure the DCT-II core

Let's analyse the core of DCT-II (10) in the matrix form:

$$X^{c2}(n) = C_N^{II}(n,m)\,x(m), \qquad (n,m = 0,1,...,N-1), \tag{12}$$

where $C^{II}{}_N$ is the basic matrix of dimension ($N \times N$) with indices on rows n = 0,1,…, N-1 and indices on columns m = 0,1…, N-1.

The function cos [n (2m + 1) π / (2N)] is periodic in relation to 4N samples, so we can write the basic matrix C^{II}_a (n,m) as integer arguments of the cosine function:

$$C_a^{II}(n,m)=[((2m+1)n \bmod 4N], \quad (n,m=0,1,...,N-1), \tag{13}$$

without taking into account $\varphi=\pi/(2N)$.

To obtain a model of the structure of the block-cyclic core DCT-II, according to [25], we analyze the matrix C^{II}_a twice as large for n, m = 0, 1,…, (2N-1) taking into account the periodicity (13),

$$C_a{}^{II}(n,m)=(c_{n,m}) \bmod (4N), \quad (n,m=0,1,...,2N-1), \tag{14}$$

where $c_{n,m}$ = (2m+1)n is the argument of the basic cosine function for each m-th element of the n-th component of DCT-II.

To form a substitution, choose two columns of the base matrix C^{II}_a. These columns are arbitrary, but should not have an index of m, which is an element of the decomposition of the size N of the transform.

The basic function cos [n(2m+1)π/(2N)] is symmetric in relation to the argument 2N, reduced according to the properties of symmetry

$$C_a^{II}(n,m)=4N-[(c_{n,m}) \bmod (4N)], \quad \text{if } [(c_{n,m}) \bmod 4N]>2N,(n,m=0,1,...2N-1). \tag{15}$$

Therefore the model of the block-cyclic structure of the basis matrix of the DCT-II can be described using a cyclic decomposition determined by the substitution of the corresponding columns in the matrix C^{II}_a (15) with integer values of the arguments of the [25]

$$H(L)=H_1(L_1)H_2(L_2)...H_k(L_k)=(h_{11},h_{12},...h_{1L1})(h_{21},h_{22},...h_{2L2})...(h_{kL1},h_{kL2},...h_{kLk}), \tag{16}$$

where h_{ij} are integer elements of cycles $H_i(L_i)$ with size L_i elements (i = 1,2,…, k; j = 1,2,…, L_i), k is the number of cycles.

The elements h_{ij} of the cycle $H_i(L_i)$ correspond to the values from the matrix of arguments of the basic harmonic transform function C^{II}_a (15) and are less than or equal to 2N. The number of cycles k in the model H(L) is determined by the specific value of the transform size N.

Due to the different expressions for the indexes of columns and rows included in the arguments of the basis function, it is necessary to use two arrays $Hr(n)$ for rows and $Hc(n/2)$ for columns to re-index the matrix C^{II}_a (5).

For example of the DCT-II of size $N = 7$, we will form a substitution on the basis of two columns C^{II}_a. The first column of arguments of the basis function in the matrix C^{II}_a with the index $n = 0$ corresponds to a natural series. The second substitution column is 3, which is the index $(2m+1)$ for $m = 1$ and is not an element of the factorization of the size N of the transform. That is, the substitution will contain elements $c_{n,m}$, which are defined according to the properties of symmetry (5),

0- column	(0 1 2 3 4 5 6 7 8 9 10 11 12 13)
1- column	(0 3 6 9 12 13 10 7 4 1 2 5 8 11) .

The substitution is described as a cyclic decomposition $H(L)$ with increasing the values of first elements hi1 for each next of cycles $H_i(L_i)$

$$H(14)=H_0(1)H_1(3)H_2(3)H_3(3)H_4(3)H_5(1) =$$

$$(0)(1,3,9)(2,6,10)(4,12,8)(5,13,11)(7),$$

where $h_{01}=0$, $h_{11}=1$, $h_{21}=2$, $h_{31}=4$, $h_{41}=5$, $h_{51}=7$.

In order to reduce the computational complexity of the transform, it is necessary to reformat the standard cyclic decomposition of substitution. For this provided that each subsequent cycle $H_{i+1}(L_{i+1})$ begins with the first element equal to $(2N-h_{i1})$, where h_{i1} is the first element of the previous cycle $H_i (L_i)$, if it exists and $L_i+1 = L_i$. In each cycle, which is formed as a result of cyclic decomposition of the substitution, the cyclic shift of the elements is possible. The cyclic decomposition $H_r(14)$ for indexing the rows of the base matrix DCT-II uses the cycles from the $H(14)$ with the corresponding first elements, which corresponds to the described condition,

$$H_r(14) = (0)(1,3,9)\ (13,11,5)\ (2,6,10)\ (12, 8, 4)\ (7).$$

The cyclic decomposition $H_c(7)$ for indexing the columns of the base matrix DCT-II uses selected cycles from $H(14)$ with the first odd elements $(2n+1)$ less than $2N$:

$$H_c(7) = (1,3,9)(13,11,5)(7),$$

the transition of the values of the elements from $(2n+1)$ to n will look like in $H_c(7)$:

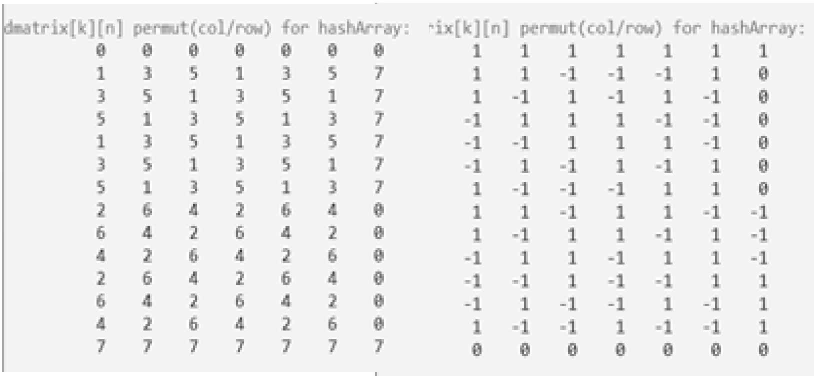
$$H_c(7) = (0,1,4)(6,5,2)(3).$$

Thus, as a result of rearrangements of sequences of rows, respectively $H_r(14)$, and columns, respectively $H_c(7)$, we obtain the basic matrix DCT-II of simplified arguments which contains a set of cyclic submatrices with integer elements.

According to the properties of asymmetric (π) of DCT-II basis function the matrix (Fig. 6) consists of the simplified elements $\underline{c_{n,m}}$ of the arguments, which are determined by the consistent arithmetic operations:

$$\underline{c_{n,m}} = 2N - \{4N - [(c_{n,m}) \bmod (4N)], \quad \text{if } \{4\mathrm{N} - [(c_{n,m}) \bmod 4N] > N, \qquad (17)$$
$$\text{otherwise} \quad \underline{c_{n,m}} = c_{n,m}, \quad (n, m = 0, 1, \ldots 2N-1).$$

The property of the symmetric/asymmetric of the basis functions of DCT-II proves efficient representation over less value of elements $\underline{c_{n,m}}$ with the addition of corresponding signs $Zc(n,m)$ in Figure 6 are denoted dmatrix and s_matrix. The matrix of signs $Zc(n,m)$ consists of the values of elements equal to +1, -1, 0 and has the same block cyclic structure with simplified arguments.



```
dmatrix[k][n] permut(col/row) for hashArray:   ix[k][n] permut(col/row) for hashArray:
      0   0   0   0   0   0   0                  1   1   1   1   1   1   1
      1   3   5   1   3   5   7                  1   1  -1  -1  -1   1   0
      3   5   1   3   5   1   7                  1  -1   1  -1   1  -1   0
      5   1   3   5   1   3   7                 -1   1   1   1  -1  -1   0
      1   3   5   1   3   5   7                 -1  -1   1   1   1  -1   0
      3   5   1   3   5   1   7                 -1   1  -1   1  -1   1   0
      5   1   3   5   1   3   7                  1  -1  -1  -1   1   1   0
      2   6   4   2   6   4   0                  1   1  -1   1   1  -1  -1
      6   4   2   6   4   2   0                  1  -1   1   1  -1   1  -1
      4   2   6   4   2   6   0                 -1   1   1  -1   1   1  -1
      2   6   4   2   6   4   0                 -1  -1   1  -1  -1   1   1
      6   4   2   6   4   2   0                 -1   1  -1  -1   1  -1   1
      4   2   6   4   2   6   0                  1  -1  -1   1  -1  -1   1
      7   7   7   7   7   7   7                  0   0   0   0   0   0   0
```

Fig. 6. The block cyclic structure with simplified arguments and the signs for size N=7

Based on the model (6) in the form of an arrays $Hr(n)$ and $Hc(n/2)$, the rows / columns of the matrix of the simplified elements $c_{n,m}$ of the arguments (17) are re-indexed, which results in the formation of block-cyclic structures C^{II}_N the core of the DCT-II.

4. The Analysis of models of block-cyclic structures of the DCT-II core

For automatic synthesis of the fast algorithm DCT-II it is necessary to perform the analysis of the structure of the obtained block-cyclic matrix C^{II}_N in order to determine identical blocks that are placed horizontally and vertically relative to each other. The presence of identical blocks reduces the computational complexity and provides an opportunity to organize the efficient computation of DCT. The criterion and initial parameters of the analysis is a search for identical blocks with the corresponding indices of placement in the basic matrix, the elements of which belong to one of the cycles of the model (16). As a result, the obtained data allow us to determine the minimum number and dimension of cyclic convolutions for further computation of DCT-II based on the fast cyclic convolutions. To research the structures of matrices a software analysis is used, which performs iterative scanning of the entire set of elements of the matrix. To find the specified fragments, a step-by-step scan of the matrix with the corresponding direction of movement is performed. The disadvantage of this scan is the large number of computations, which with increasing dimension of the matrix of size $N \times N$ has the order of computational complexity $O(N^4)$. The values of the elements of the matrix $N \times N$ can be determined in advance, but the large sizes of transforms require significant memory costs to save them. For automatic synthesis of algorithms for computing DCT-II based on cyclic convolutions, it is necessary to provide efficient, in terms of speed, analysis of the structure of the block-cyclic core of transform. To do this, we apply the basic parameters of the model $H(L)$ of the block-cyclic structure of the basic matrix DCT-II. The model $H(L)$ of the block-cyclic structure of the base matrix DCT-II complements its representation in the form of a simplified cyclic decomposition $H'(L)$, the simplified elements h'_{ij} of which are determined by formula (17), and additional cyclic decomposition of signs $Zc(L)$

$$H'(L) = H'_1(L_1)H'_2(L_2)...H'_k(L_k) = \\ = (h'_{11}, h'_{12}, ...h'_{1L1})(h'_{21}, h'_{22}, ...h'_{2L2})...(h'_{kL1}, h'_{kL2}, ...h'_{kLk}), \quad (18)$$

$$Zc(L) = Z_1(L_1)\, Z_2(L_2) \ldots Z_k(L_k) = \tag{19}$$
$$= (z_{11}, z_{12}, \ldots z_{1L1})\,(z_{21}, z_{22}, \ldots z_{2L2}) \ldots (z_{kL1}, z_{kL2}, \ldots z_{kLk}),$$

where z_{ij} =1, if $h_{ij} < N$, or z_{ij} =-1, if $N < h_{ij} \leq 2N$, or h_{ij} =0 if $h_{ij} = N$; *i=1,2,…,k; j=1,2,…,Li.*

Therefore, accordance the model (8, 9) for the analysis of the block-cyclic structure of the core of the DCT-II, we use the parameters:

k is the number of cycles in *H*(*L*);

Li is the number of elements in each cycle *Hi*(*Li*);

t is the repetition of groups of elements h'_{ij} of the corresponding cycle *Hi'*(*Li*) in the row of submatrices of the matrix structure;

i,j is coordinates of the first elements of submatrices with the corresponding sign and value of $z_{ij}c_{ij}$ in the matrix structure;

r is the number of repetitions of identical cyclic submatrices in the matrix structure horizontally/vertically;

r' is the number of submatrices, starting not from the first $c_{ij} = h'_{i1}$ but from the intermediate simplified element h'_{ij} of the corresponding cycle *Hi'*(*Li*);

m is the total number of submatrices in the block-cyclic structure.

The block-cyclic structure of a square matrix C_a (*k*,*n*) contains a set of cyclic submatrices of different size *Li*. These submatrices contain integers and are Latin squares. In addition, each square submatrix contains equal elements arranged parallel to the side diagonal or equal pairs of elements symmetrically arranged relative to the main diagonal. Such square submatrices are called Hankel or left-circulant, completely defined by their first row or first column. The model of the block-cyclic structure of the core of the DCT-II of size *N*, which contains left-circulant submatrices, is determined by cyclic decompositions of substitution (18, 19) with the corresponding parameters. Therefore, to find and determine identical submatrices in the basic matrix, it is advisable to use an algorithm based on the parameters of cyclic decompositions. Taking into account the peculiarities of submatrices (square, left-circulant) in the block-cyclic structure of the basic matrix DCT-II is possible to analyse identical submatrices by different algorithms in the directions of search and comparison. The search and analysis by sorting through the elements, starting from the top row or the first column, involves checking for cyclicity (the same values of elements with the next offset row or column). In the case of non-fulfillment of the requirement for cyclicity, we conclude that the size of the submatrix is complete, and the obtained size determines the dimension of the square submatrix. Next, we move from top to bottom and left to right (horizontally / vertically) to this defined size. We continue the analysis by searching the elements of the first rows or columns, determining a new submatrix, compare it with previously identified submatrices for identity. The search and analysis by sorting through the elements on the lateral diagonal begins with the first element by checking for cyclicity (the same values of elements parallel to the side diagonal). If the requirement of cyclicity is not met or the maximum value of the number of elements in the lateral diagonal is determined, we conclude that the size of the submatrix is complete and determine its dimension. Next, move from top to bottom and left to right (horizontally/vertically) to this defined size. Continue the analysis by searching for elements in the lateral diagonal when determining a new submatrix, compare it with previously identified, determining the same separately for the same vertical and horizontal coordinates. Mixed search for the values of elements in the row / column and the values of elements relative to the lateral diagonal, combining these two strategies, taking into account the peculiarities of the corresponding sizes of submatrices will also identify identical cyclic submatrices in the block-cyclic structure of the basic matrix DCT-II. Identical cyclic submatrices are identical submatrices that have the same corresponding values of elements from the simplified cycle $Hi'(Li)$ and the sign cycle $Zc(L_i)$. Quasi-identical cyclic submatrices have the same elements from the simplified cycle $Hi'(Li)$, but opposite values in the cyclic decomposition of signs $Zc(L_i)$. An important feature of the block-cyclic structure of the formed basic matrix DCT-II is presence of square, left-circulant integer submatrices, the location coordinates, elements and dimensions of which are determined by the corresponding cycles in (8, 9). This feature allows you to speed up the analysis of the structure of the matrix by changing the scanning step equal to the size of the cycle L_i, containing the corresponding integer elements h_{ij} in the cycle *Hi*(*Li*). The analysis of the matrix structure for the identity of the submatrices placed in it will be performed by coordinates (*i*, *j*), for which we will determine the values of $z_{ij}\underline{c_{i,j}}$ of the first elements of submatrices. Next, it is compared ($c_{i,j}$) with the corresponding values h'_{ij} of the first or other elements of the cycle *Hi'*(*Li*) and as a result of the correspondence, the dimension *Li* × *Li* of the square submatrix is determined. Further selection of identical horizontally and vertically placed matrices under the condition of equality of their first elements is performed by the coordinates of the column / row at offsets on the size *Li*. The following coordinates (Table 1) of the first elements of submatrices are determined by (*i*+*Li*), (*j*+ *Li*), where the dimension *Li* is chosen according to the values of the first elements of submatrices in the matrix structure for the cycle *Hi'*(*Li*).

Table 4

The matrix structure with coordinates and values $z_{i,j}\ \underline{c}_{i,j}$ of the first elements with the sign for cyclic submatrices of the DCT-II

(1,1) – $z_{i,j}\ \underline{c}_{i,j}$;		**(1,1+ L_1) – $z_{i,j}\ \underline{c}_{i,j}$;**		**…**	**(1, 1+ L_1+L_2+…+L_k) – $z_{i,j}\ \underline{c}_{i,j}$;** **(1+ L_k, 1+ L_1+ L_2 + …+L_k) – $z_{i,j}\ \underline{c}_{i,j}$;**
(1+ L_1,1) – $z_{i,j}\ \underline{c}_{i,j}$;		**(1+ L_1,1+ L_1) – $z_{i,j}\ \underline{c}_{i,j}$;**		**…**	**(1+2L_k, 1+L_1+L_2+…+L_k) – $z_{i,j}\ \underline{c}_{i,j}$;** **(1+ 3$L_k$, 1+ L_1+ L_2+…+L_k) – $z_{i,j}\ \underline{c}_{i,j}$;**
(1+ L_1+ L_2,1) – $z_{i,j}\ \underline{c}_{i,j}$;		**(1+ L_1+ L_2,1+ L_1) – $z_{i,j}\ \underline{c}_{i,j}$;**		**…**	**(1+ 4L_k, 1+ L_1+ L_2+…+L_k) – $z_{i,j}\ \underline{c}_{i,j}$;**
					…
(1+L_1+L_2+…+L_k,1) – $z_{i,j}\ \underline{c}_{i,j}$;	**(1+L_1+L_2+…+L_k, 1+L_k) – $z_{i,j}\ \underline{c}_{i,j}$;**	**(1+L_1+L_2+ … +L_k,1+2L_k) – $z_{i,j}\ \underline{c}_{i,j}$;**	**(1+L_1+L_2+ … +L_k,1+3L_k) – $s_{i,j}\ \underline{c}_{i,j}$;**		**(1+L_1+ L_2 + …+L_k, 1+L_1+ L_2+…+L_k) – $z_{i,j}\ \underline{c}_{i,j}$;**

We define identical cyclic submatrices by selecting the coordinates ($i+L_i$), ($j+L_i$) of the first elements $z_{i,j}\underline{c}_{i,j}$ of identical submatrices horizontally ($i+L_i$) = const in the block-cyclic structure of the basis matrix. For horizontally identical cyclic submatrices, cyclic convolution with element-wise addition/subtraction input data will be performed. Similarly, identical cyclic submatrices are determined by selecting the coordinates ($i+L_i$), ($j+L_i$) of the first elements of identical submatrices vertically ($j+Li$) = const in the block-cyclic structure of the basis matrix. For vertically placed identical cyclic submatrices, one cyclic convolution will be performed with the corresponding input data.

For the remaining cyclic submatrices, that do not have identical block-cyclic structure of core of DCT-II in size N, one cyclic convolution will be performed.

5. The software implementation of the synthesis of fast DCT-II algorithms based on cyclic convolutions

The software implementation of the synthesis of fast DCT-II algorithms using the cyclic substitutions technique includes the analysis and search of identical matrices.

It includes two main functions: search and selection of cyclic matrices in the block-cyclic structure of the core of DCT-II based on the model (18, 19) and a function of determination of identical blocks for the definition of the minimum number of cyclic convolutions required for computation of DCT-II.

The software solution is implemented using C++ in the development environment of Visual Studio C++ 2022.

The first function searches for and determines the affiliation of cyclic blocks to the corresponding cycles $Hi'(Li)$ with the same first elements $\underline{c}_{i,j}$.

The analysis and selection of cyclic blocks is performed for each cycle $Hi'(Li)$ of the model of the block-cyclic structure of the DCT-II of size N.

The analysis is the research of the obtained data set with values of $\underline{c}_{i,j}$ of the first elements of cyclic blocks and their coordinates (i, j) in the structure of the block-cyclic matrix of the simplified elements $\underline{c}_{n,m}$.

The analysis is aimed at identifying identical blocks placed horizontally and vertically relative to each other. The block diagram of the algorithm for determining identical blocks and, accordingly, the minimum number of cyclic convolutions is given in Figure 7.

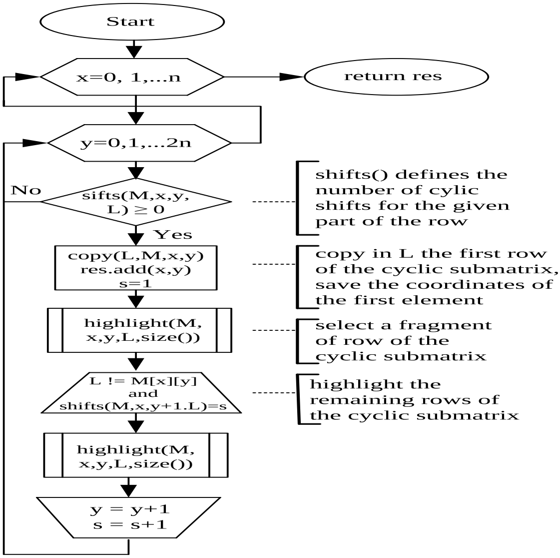


*Fig. **7.** Flowchart of the algorithm for determining identical blocks*

Table 2

The values $\underline{\mathbf{c}}_{i,j}$ of the first elements of the blocks and their coordinates (i, j)

$\underline{c}_{i,j}$	(*i*, *j*) – coordinates of the first elements of the blocks
0	(0, 0), (13, 12), (14, 12), (15, 12), (16, 12), (17, 12), (18,12);
1	(1, 0), (4, 0), (1, 3), (4, 3);
3	(19, 6), (22, 6), (19, 9), (22, 9);
5	(19, 0), (22, 0), (19, 3), (22, 3), (1, 6), (4, 6), (1, 9), (4, 9);
2	(7, 0), (10, 0), (7, 3), (10, 3);
10	(7, 0), (10, 0), (7, 3), (10, 3);
4	(13, 0), (16, 0), (13, 3), (16, 3);
8	(13, 6), (16, 6), (13, 9), (16, 9);
7	(25, 0), (26, 0), (1, 12), (2, 12), (3, 12), (4, 12), (5, 12), (6, 12), (19, 12), (20, 12), (21, 12), (22, 12), (23, 12), (24, 12).

As a result of execution of the first function, a set of cyclic subarrays is determined in the form of an array of data containing the values of $\underline{c}_{i,j}$ of the first simplified elements of the blocks and their corresponding coordinates

(i,j), which accordingly are the indices of the two-dimensional array. For example, as a result of the analysis for DCT-II for size N = 14, we obtain an array of data shown in Table 2. According to Table 2, the first elements of cyclic subarrays of the basis matrix, equal to, for example, $\underline{c}_{i,j}$ = 1, are located in the matrix at coordinates 1 (1,0); 1 (4.0); 1 (1,3); 1 (4.3). That is, in the structure of the basic matrix there are two identical blocks, placed horizontally (1 (1.0); 1 (1.3) and 1 (4.0); 1 (4.3)), and vertically (1 (1) , 0); 1 (4,0) and 1 (1,3); 1 (4,3)) relative to each other. For clarity of the results of the analysis in Table 2, in Figure 8, we show the block-cyclic structure of the arguments of the basic matrix DCT-II of size N = 14 with the coloring of the first elements and identical cyclic matrices, which in the analysis are not calculated, except for integer values of the first elements of submatrices.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	3	9	1	3	9	5	13	11	5	13	11	7	7
2	3	9	1	3	9	1	13	11	5	13	11	5	7	7
3	9	1	3	9	1	3	11	5	13	11	5	13	7	7
4	1	3	9	1	3	9	5	13	11	5	13	11	7	7
5	3	9	1	3	9	1	13	11	5	13	11	5	7	7
6	9	1	3	9	1	3	11	5	13	11	5	13	7	7
7	2	6	10	2	6	10	10	2	6	10	2	6	14	14
8	6	10	2	6	10	2	2	6	10	2	6	10	14	14
9	10	2	6	10	2	6	6	10	2	6	10	2	14	14
10	2	6	10	2	6	10	10	2	6	10	2	6	14	14
11	6	10	2	6	10	2	2	6	10	2	6	10	14	14
12	10	2	6	10	2	6	6	10	2	6	10	2	14	14
13	4	12	8	4	12	8	8	4	12	8	4	12	0	0
14	12	8	4	12	8	4	4	12	8	4	12	8	0	0
15	8	4	12	8	4	12	12	8	4	12	8	4	0	0
16	4	12	8	4	12	8	8	4	12	8	4	12	0	0
17	12	8	4	12	8	4	4	12	8	4	12	8	0	0
18	8	4	12	8	4	12	12	8	4	12	8	4	0	0
19	5	13	11	5	13	11	3	9	1	3	9	1	7	7
20	13	11	5	13	11	5	9	1	3	9	1	3	7	7
21	11	5	13	11	5	13	1	3	9	1	3	9	7	7
22	5	13	11	5	13	11	3	9	1	3	9	1	7	7
23	13	11	5	13	11	5	9	1	3	9	1	3	7	7
24	11	5	13	11	5	13	1	3	9	1	3	9	7	7
25	7	7	7	7	7	7	7	7	7	7	7	7	7	7
26	7	7	7	7	7	7	7	7	7	7	7	7	7	7
27	14	14	14	14	14	14	14	14	14	14	14	14	14	14

Fig. 8. The matrix of DCT-II of size N = 14 with the found identical cyclic blocks

The algorithm for determining the minimum number of convolutions uses the result of the first search and selection of cyclic submatrices. In the next stage of the synthesis, the values of the first simplified elements of cyclic submatrices are selected and are compared in the structure of the basis matrix according to the formed coordinates. For each group of coordinates of the found first elements $\underline{c}_{i,j}$ on the same coordinate y_i (i = 0,1,…, 2N-1), a set of convolutions is formed using the function *get_conv*(*v*). After each iteration, the processed convolutions are added to the unique list. As a result of definition of identical blocks horizontally and vertically by means of the software decision, the sequence of cyclic convolutions between a set of the simplified arguments and a set of the corresponding input data is formed.

For example, for DCT-II of size $N = 14$, we obtain (Fig. 9) a conditionally written sequence of cyclic convolutions (X) for the corresponding values of the first elements with coordinates $z_{i,j}\underline{c}_{i,j}$ $\{ i, j \}$. To compute DCT-II of size $N = 14$ according to the formed block-cyclic structure (Fig. 8) it is necessary to perform 4 one-point and 8 three-point cyclic convolutions. Convolution number 5.c was determined to be opposite in sign to convolution number 2.c based on the analysis of identical cyclic submatrices vertically. At this stage of the analysis, the equals but opposite in sign of the convolutions are considered to be of the same type.

As a result of the analysis of the structure of the formed block-cyclic matrix for the identity of the blocks placed horizontally and vertically, we determine the reduced number and size of cyclic convolutions for the synthesis of the computation algorithm of DCT-II. A set of a reduced number of identical cyclic submatrices in the structure of the basic matrix determines the organization of the computations DCT-II of size *N*.

```
Convolutions:
1.   a)  0 { 0,  0}   ->  ( +0) (X) { +x(0)  +x(1)  +x(4)  +x(13)  +x(12)  +x(9)  +x(2)  +x(7)  +x(5)  +x(11)  +x(6)  +x(8)  +x(3)  +x(10)  }

2.   a) +1 { 1,  0} -1 { 1,  3}   ->  ( +1  +3  +9) (X) { +(x(0), x(1), x(4))  -(x(13), x(12), x(9))  }
     b) +5 { 1,  6} -5 { 1,  9}   ->  ( +5  -13  +11) (X) { +(x(2), x(7), x(5))  -(x(11), x(6), x(8))  }
     c) +7 { 1, 12}   ->  ( +7) (X) { +x(3)  -x(10)  }

3.   a) +2 { 7,  0} +2 { 7,  3}   ->  ( +2  +6  -10) (X) { +(x(0), x(1), x(4))  +(x(13), x(12), x(9))  }
     b) +10 { 7,  6} +10 { 7,  9}   ->  ( +10  -2  -6) (X) { +(x(2), x(7), x(5))  +(x(11), x(6), x(8))  }

4.   a) +4 {13,  0} +4 {13,  3}   ->  ( +4  +12  -8) (X) { +(x(0), x(1), x(4))  +(x(13), x(12), x(9))  }
     b) -8 {13,  6} -8 {13,  9}   ->  ( -8  +4  +12) (X) { -(x(2), x(7), x(5))  -(x(11), x(6), x(8))  }
     c)  0 {13, 12}   ->  ( -0) (X) { -x(3)  -x(10)  }

5.   a) +5 {19,  0} -5 {19,  3}   ->  ( +5  -13  +11) (X) { +(x(0), x(1), x(4))  -(x(13), x(12), x(9))  }
     b) -3 {19,  6} +3 {19,  9}   ->  ( -3  -9  +1) (X) { -(x(2), x(7), x(5))  +(x(11), x(6), x(8))  }
     c) -7 {19, 12}   ->  ( -7) (X) { -x(3)  +x(10)  }      // repeated (opposite)

6.   a) +7 {25,  0}   ->  ( +7) (X) { +x(0)  -x(1)  +x(4)  -x(13)  +x(12)  -x(9)  -x(2)  +x(7)  -x(5)  +x(11)  -x(6)  +x(8)  +x(3)  -x(10)  }
```

Fig. 9. The formed sequence of cyclic convolutions for DCT-II of size N = 14

6. Discussion of the results of analysis of the models the block-cyclic structure of DCT-II cores

The formation of the basic matrix of DCT in the form of a set of cyclic submatrices and analysis of block structure with variable step allows us to speed up the process of analysis of the structure of the core of the transform, reduce the number of computations of cyclic convolutions and, consequently, reduce computational complexity. The reduced number of cyclic submatrices in the structure of the basis matrix depends on the value of the transform size N and model (18, 19) of the block-cyclic structure of the core of DCT-II of size N. Identical cyclic submatrices placed vertically relative to each other lead to a one-time computation of cyclic convolutions, the results of which are used in the process of combining for different $X^{2c}(k)$ output values of the transform (1). Identical cyclic submatrices placed horizontally relative to each other lead to the union of groups of input values $x(n)$ of the transform and a one-time computation of cyclic convolutions, the results of which are used for one group $X^{2c}(k)$ of output values of the transform (1). Depending on the specific size of the transform N, we have in the block-cyclic structures a corresponding number of cyclic convolutions, which is not constantly increasing with an increasing size N of transform. In the process of analysis the block-cyclic structures of the DCT-II core for sizes $N = p^i$, the regular increase of horizontal and vertical lines in the structure of transform basis with increasing size of degree is confirmed, which is shown in Fig. 5 for sizes $N = 3^i$ on the example of sizes of $N =$ 9,27. As it can be seen in the example for $N = 3^i$ (Fig. 10), for block-cyclic structures of the core of DCT-II for sizes of the integer $N = p^i$ is characterized by a regular increase of horizontal and vertical lines in the structure of the basis. Similarly, this is confirmed by the corresponding increase in the number of cycles in the model (18, 19) of the block-cyclic structure of the core of the DCT-II with an increasing value of the degree of size $N = p^i$. Depending on the specific value of the transform of the size N according to the corresponding value, the choice of columns to form a substitution for the model (18, 19) of the block-cyclic structure of the core of the DCT-II gives different variants of block-cyclic structures with the corresponding number and sizes of cyclic matrices. For example, in Table 3 for DCT-II of size *N*=20 according to the values of the indices of columns 1, 3, 4, 5, 9, we have four variants of the block-cyclic structures

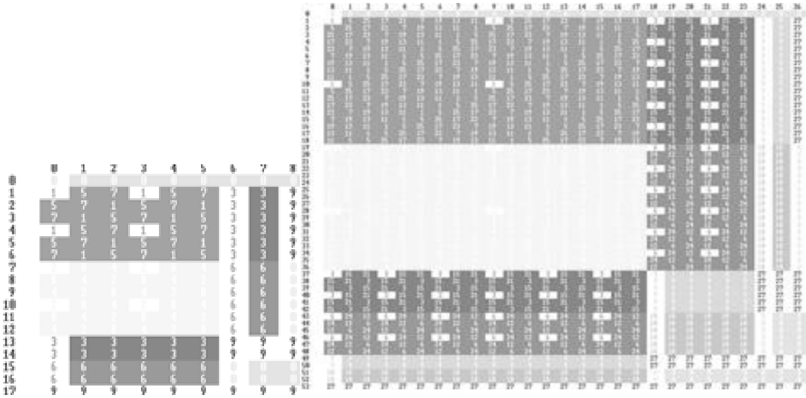


Fig. 10. The block-cyclic structures of the basic matrix of arguments DCT-II sizes N = 9, 27

Table 3

The values point and number of cyclic convolutions for DCT-II of size N=20 corresponding to the value of column to form a substitution

N	Col	point (number)
20	1	1(4), 2(7), 4(6)
	3	1(12), 2(4), 4(6)
	4	1(20), 2(24)
	5	1(12), 2(28)
	9	1(28), 2(20)
	19	1(110)

In the example for DCT-II of size *N*=20 according to choice the values of the indices of columns 1, we obtain a conditionally written sequence of cyclic convolutions (X) between the arguments of the basic function and the input sequence (Fig. 11).

To compute DCT-II of size *N*=20 based on the formed block-cyclic structure (Fig. 11), it is necessary to perform 4 one-, 7 two- and 6 four-point cyclic convolutions accordance Table 3. Convolution number 6.b was determined to be the opposite of convolution number 2.b based on the analysis of identical cyclic submatrices in the vertical direction. This creates the possibility of choosing the structural diagrams of computers of DCT-II at the system (algorithmic) stage of design. There are sizes of DCT-II, which have only one variant of the block-cyclic structure for any values of column indices.

For example, one variant of the block-cyclic structure has simple values of sizes $N = 11, 23, 47, 59, 83$.

Computations of DCT-II for short sizes based on cyclic convolutions are characterized by the least time among other approaches to computing transforms [26] and reducing the cost of implementing computers on VLSI [14, 15].

The synthesized DCT-II algorithms as a result of the analysis of the block-cyclic structures for short sizes are basic in the transition to large composite sizes of transform.

```
1. a)  0 { 0,  0}   -> ( +0) (X) { +x(0)  +x(1)  +x(4)  +x(13)  +x(19)  +x(18)  +x(15)  +x(6)  +x(2)  +x(7)  +x(17)  +x(12)  +x(3)  +x(10)  +x(8)
       +x(14)  +x(16)  +x(9)  +x(11)  +x(5)  }

2. a) +1 { 1,  0} -1 { 1,  4}   -> ( +1  +3  +9  -13) (X) { +(x(0), x(1), x(4), x(13)) -(x(19), x(18), x(15), x(6)) }
   b) +5 { 1,  8} -5 { 1, 10}   -> ( +5  +15) (X) { +(x(2), x(7)) -(x(17), x(12)) }
   c) +7 { 1, 12} -7 { 1, 16}   -> ( +7  -19  +17  -11) (X) { +(x(3), x(10), x(8), x(14)) -(x(16), x(9), x(11), x(5)) }

3. a) +2 { 9,  0} +2 { 9,  4}   -> ( +2  +6  +18  -14) (X) { +(x(0), x(1), x(4), x(13)) +(x(19), x(18), x(15), x(6)) }
   b) +10 { 9,  8}   -> ( +10) (X) { +x(2)  -x(7)  +x(17)  -x(12)  }
   c) +14 { 9, 12} +14 { 9, 16}   -> ( +14  -2  -6  -18) (X) { +(x(3), x(10), x(8), x(14)) +(x(16), x(9), x(11), x(5)) }

4. a) +4 {17,  0} -4 {17,  2} +4 {17,  4} -4 {17,  6}   -> ( +4  +12) (X) { +(x(0), x(1)) -(x(4), x(13)) +(x(19), x(18)) -(x(15), x(6)) }
   b) -12 {17, 12} +12 {17, 14} -12 {17, 16} +12 {17, 18}   -> ( -12  +4) (X) { -(x(3), x(10)) +(x(8), x(14)) -(x(16), x(9)) +(x(11), x(5)) }

5. a) +5 {21,  0} -5 {21,  2} -5 {21,  4} +5 {21,  6} -5 {21, 12} +5 {21, 14} +5 {21, 16} -5 {21, 18}   -> ( +5  +15) (X) { +(x(0), x(1)) -(x(4),
       x(13)) -(x(19), x(18)) +(x(15), x(6)) -(x(3), x(10)) +(x(8), x(14)) +(x(16), x(9)) -(x(11), x(5)) }
   b) -15 {21,  8} +15 {21, 10}   -> ( -15  +5) (X) { -(x(2), x(7)) +(x(17), x(12)) }

6. a) +7 {25,  0} -7 {25,  4}   -> ( +7  -19  +17  -11) (X) { +(x(0), x(1), x(4), x(13)) -(x(19), x(18), x(15), x(6)) }
   b) -5 {25,  8} +5 {25, 10}   -> ( -5  -15) (X) { -(x(2), x(7)) +(x(17), x(12)) }
   c) -9 {25, 12} +9 {25, 16}   -> ( -9  +13  -1  -3) (X) { -(x(3), x(10), x(8), x(14)) +(x(16), x(9), x(11), x(5)) }

7. a) +8 {33,  0} +8 {33,  2} +8 {33,  4} +8 {33,  6}   -> ( +8  -16) (X) { +(x(0), x(1)) +(x(4), x(13)) +(x(19), x(18)) +(x(15), x(6)) }
   b)  0 {33,  8}   -> ( -0) (X) { -x(2)  -x(7)  -x(17)  -x(12)  }
   c) -16 {33, 12} -16 {33, 14} -16 {33, 16} -16 {33, 18}   -> ( -16  +8) (X) { -(x(3), x(10)) -(x(8), x(14)) -(x(16), x(9)) -(x(11), x(5)) }

8. a) +10 {37,  0}   -> ( +10) (X) { +x(0)  -x(1)  +x(4)  -x(13)  +x(19)  -x(18)  +x(15)  -x(6)  -x(2)  +x(7)  -x(17)  +x(12)  +x(3)  -x(10)  +x(8)
      -x(14)  +x(16)  -x(9)  +x(11)  -x(5)  }
```

Fig. 11. Formed sequence of cyclic convolutions for DCT-II of size N=20 according to choice the values of the indices of columns 1

Thus, the analysis of the models of the block-cyclic structures of the core of the DCT-II allows us, in the process of synthesis of algorithms, to provide efficient software or hardware organization of transforms based on cyclic convolutions for each specific size *N*.

6. Conclusions with perspectives

The synthesis of fast DCT-II algorithms using the cyclic substitutions technique including the analysis of model of the block-cyclic structures of the DCT-II is obtained in the work. The analysis of the block-cyclic structure of the basic matrix for identical blocks is performed using the variable search step, based on the models of the block-cyclic structures of the core of the DCT-II. Algorithmic support and software for analysis of the structure of the block-cyclic basis matrix have been developed, which is used to determine an array of parameters for the formal description of the structure of core of the DCT-II. As a result of the analysis of the models of the block-cyclic structures of core of the DCT-II, a reduced number of identical cyclic submatrices is determined, which allows us to reduce the number of cyclic convolutions for computing of the DCT-II of arbitrary sizes N. A fast cosine transform algorithm implements in the form of a set of operations of cyclic convolutions over combined sequences of input data and coefficients of the basic function of transform.

The practical significance of the work lies in the fact that the obtained results of synthesis of fast DCT-II algorithms using the cyclic substitutions technique for specific sizes of transform is important for the system engineering stage of designing a DCT-II based on cyclic convolutions, because clearly reflect the quantitative interactions of its parts on algorithmic level. These transforms are used in information technologies for various purposes, especially in convolutional neural networks [27].

Prospects for further research are the development the synthesis of fast algorithms of the DCT-II in parallel implementation and analysis the block-cyclic structures of the DCT-II core for large sizes.

7. References

[1] C.S. Burrus, M. Frigo, S.G. Johnson, M. Pueschel, I. Selesnick, Fast Fourier Transforms, Samurai Media Limited, 2018.

[2] IEEE Standard for Second-Generation IEEE 1857 Video Coding. IEEE STD 1857.4-2018. Institute of Electrical and Electronics Engineers. August 30, 2019. pp. 1–199. DOI:10.1109/IEEESTD.2019.8821610.

[3] L. Stankovic, Digital signal processing. Basic Theory and Applications. Revised edition, North Charleston, South Carolina, USA, 2020.

[4]V. Britanak, K.R. Rao, Cosine-/Sine-Modulated Filter Banks. Springer International Publishing. 2018. DOI:10.1007/978-3-319-61080-1

[5]A. Cariow, L. Lesiecki, Small-size algorithms for type-IV discrete cosine transform with reduced multiplicative complexity. Radioelectronics and Communications Systems 63 9 (2020) 465–487. DOI:10.3103/S0735272720090022

[6]G. Malaschonok, A. Ivashkevich, Quick Recursive QR Decomposition, in: Proceedings of the Conference on Mathematical Foundations of Informatics (MFOI2020), 12-16 January 2021, Kyiv, Ukraine, pp. 268-275.

[7] I. Tsmots, V. Rabyk, N. Kryvinska, M. Yatsymirskyy, V. Teslyuk, Design of the Processors for Fast Cosine and Sine Fourier Transforms, Circuits, Systems, and Signal Processing, 41 9 (2022) 4928-4951. DOI:10.1007/s00034-022-02012-8

[8] O. Beaumont, Ph. Duchon, L. Eyraud-Dubois, J. Langou, M. Verite, Symmetric block-cyclic distribution: fewer communications leads to faster dense cholesky factorization, in: Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis (SC '22), Dallas, Texas, 2022, pp. 1–15.

[9] A. Kaveh, Optimal analysis of structures by concepts of symmetry and regularity", 1st ed., Springer, Vienna, Austria, 2013. DOI:10.1007/978-3-7091-1565-7

[10]C. M. Rader, Discrete Fourier transform when the number of data samples is prime, in: Proceedings IEEE, 1968, vol. 56, pp. 1107-1108.

[11]R.E. Blahut, Fast Algorithms for Signal Processing, Cambridge University Press, Cambridge, 2010.

[12]M. Teixeira, Y.I. Rodríguez, Parallel cyclic convolution based on recursive formulations of block pseudocirculant matrices, in: IEEE Trans. Signal Processing, 56 7 (2008) 2755-2770.

[13]A.N. Tereshchenko, Optimization of Pitassi method for calculating convolution, Artificial Intelligence 1 (2009) 204–212.

[14]P.K, Meher, S.K. Lam, T. Srikanthan, D.H. Kim, S.Y. Park, Area-Time Efficient Two-Dimensional Reconfigurable Integer DCT Architecture for HEVC, Electronics 10 5, 603 (2021). DOI:10.3390/electronics10050603.

[15]D.F. Chiper, L.-T. Cotorobai, An Efficient Algorithm for the VLSI Implementation of Inverse DCT Based on Quasi-Circular Correlation Structures. in: 15th International Conference on Advanced Technologies, Systems and Services in Telecommunications (TELSIKS), Nis, Serbia, 20-22 Oct. 2021, pp.197–201. DOI:10.1109/TELSIKS52058.2021.9606373.

[16]G. Goertzel, An algorithm for the evaluation of finite trigonometric series, Amer. Math. Monthly, 65 (1958) 34-35.

[17] L.I. Bluestein, Linear filtering approach to the computation of discrete Fourier transforms, in: IEEE Trans. Audio Electroacoust., AU-18 (1970) 451-455.

[18]S. Winograd, On computing the discrete Fourier transforms, Mathematics of Computation, 32 (1978) 175-199.

[19]D. P. V. Muddhasani, M. D. Wagh, Bilinear algorithms for discrete cosine transforms of prime lengths, Signal Processing, 86 9 (2006) 2393-2406.

[20]P. K. Meher, Systolic designs for DCT using a low complexity concurrent convolutional formulation, in: IEEE Trans. Circuits & Systems for Video Technology, 16 10 (2006) 1041–1050.

[21] M. Puschel, J. M. F. Moura, The Algebraic Approach to the Discrete Cosine and Sine Transforms and their Fast Algorithms, SIAM Journal of Computing, 32.5 (200): 1280-1316.

[22]I. Prots'ko, Enhancement of the Efficient Computation of Real Harmonic Transforms Based on Cyclic Convolutions. Thesis for a scientific degree of Doctor of Engineering, Vinnytsia National Technical University, Vinnytsia. 2019.

[23]S. A. Martucci, Symmetric convolution and discrete sine and cosine transform. IEEE Trans. on Signal Processing, 42 (1994)1038-1051.

[24]A.V. Oppenheimer, R.W. Schafer, Discrete-Time Signal Processing. 3rd ed., Englewood Cliffs, NJ: Prentice Hall. Pearson Education Limited, 2014.
[25]I. Prots'ko, M. Mishcuk, Block-Cyclic Structuring of the Basis of Fourier Transforms Based on Cyclic Substitution, Cybernetics and Systems Analysis, 57 6 (2021) 1008–1016.
[26]I. Prots'ko, R. Rykmas, The Runtime Benchmarking of DCT-II based on Cyclic Convolutions. International Journal of Condition Monitoring and Diagnostic Engineering Management 21 2 (2018) 11-16.
[27]S. Ju, Y. Lee, S. Lee, Convolutional Neural Networks with Discrete Cosine transform features. IEEE Transactions on Computers 71 12 (2022) 3389 – 3395. DOI: 10.1109/TC.2022.3150574

УДК 681.391

СИНТЕЗ ШВИДКИХ АЛГОРИТМІВ DCT-II ВИКОРИСТАННЯ ТЕХНІКИ ЦИКЛІЧНИХ ЗАМІН

Dc.Sci. Процько І., Рикмас Р.
Національний університет «Львівська політехніка», Україна
E-mail: ihor.o.protsko@lpnu.ua, roman.rikmas@gmail.com

Анотація. *Підхід до синтезу дискретних косинусних перетворень на основі циклічних згорток вимагає формування блочно-циклічної структури базової матриці, аналізу блочно-циклічної структури ядра перетворення та визначення мінімальної кількості циклічних згорток. . Розглянуто модель стисненого опису блочно-циклічної структури ядра дискретного косинусного перетворення у вигляді циклічної декомпозиції підстановки. Циклічний розклад підстановки містить цілі елементи аргументів базової функції перетворення. Крім того, моделі доповнюються циклічною декомпозицією підстановки зі спрощеними значеннями елементів аргументів базової функції та циклічною декомпозицією підстановки знаків. Аналіз та визначення ідентичних циклічних підматриць у базовій матриці перетворення виконується зі змінним кроком пошуку на основі параметрів моделі дискретного косинусного перетворення. В результаті програмна реалізація з використанням техніки циклічних підстановок дозволяє прискорити процес аналізу блочно-циклічної структури ядра перетворення, зменшити кількість циклічних згорток для обчислення дискретних косинусних перетворень довільних розмірів.*

Ключові слова: *дискретні косинусні перетворення; синтез алгоритму; аналіз блочно-циклічної структури; циклічне розкладання заміщення*

Наукове видання

ОБРОБКА ІНФОРМАЦІЇ В СИСТЕМАХ УПРАВЛІННЯ ТА ПРИЙНЯТТЯ РІШЕНЬ. ПРОБЛЕМИ ТА РІШЕННЯ

Монографія

Під наук. ред. проф. В. Вичужаніна

Укр. та англ. мовами

Підписано до друку 27.09.2023
Формат 60х84/16. Папір офсетний. Ум. друк. арк. 20,88
Тираж 300 пр. Замовлення № И23-10-42

НУ «ОМА», центр «Видавінформ»
65052, м. Одеса, Дідріхсона, 8, корп. 7
Свідоцтво ДК № 1292 від 20.03.2003
e-mail: publish@onma.edu.ua
Телефони: +38 (048) 793-24-50
+38 (048) 793-24-51