ISSN 2500-316X (Online)

RUSSIAN TECHNOLOGICAL JOURNAL

РОССИЙСКИЙ ТЕХНОЛОГИЧЕСКИЙ ЖУРНАЛ

Information systems.
Computer sciences.
Issues of information security

Multiple robots (robotic centers) and systems.
Remote sensing and nondestructive testing

Modern radio engineering and telecommunication systems

Micro- and nanoelectronics.
Condensed matter physics

Analytical instrument engineering and technology

Mathematical modeling

Economics of knowledge-intensive and high-tech enterprises and industries.
Management in organizational systems

Product quality management. Standardization

Philosophical foundations of technology and society

12+
www.rtj-mirea.ru

13(5) 2025

ISSN 2500-316X (Online)

RUSSIAN TECHNOLOGICAL JOURNAL

РОССИЙСКИЙ ТЕХНОЛОГИЧЕСКИЙ ЖУРНАЛ

- Information systems. Computer sciences. Issues of information security
- Multiple robots (robotic centers) and systems. Remote sensing and nondestructive testing
- Modern radio engineering and telecommunication systems
- Micro- and nanoelectronics. Condensed matter physics
- Analytical instrument engineering and technology
- Mathematical modeling
- Economics of knowledge-intensive and high-tech enterprises and industries. Management in organizational systems
- Product quality management. Standardization
- Philosophical foundations of technology and society

- Информационные системы. Информатика. Проблемы информационной безопасности
- Роботизированные комплексы и системы. Технологии дистанционного зондирования и неразрушающего контроля
- Современные радиотехнические и телекоммуникационные системы
- Микро- и наноэлектроника. Физика конденсированного состояния
- Аналитическое приборостроение и технологии
- Математическое моделирование
- Экономика наукоемких и высокотехнологичных предприятий и производств. Управление в организационных системах
- Управление качеством продукции. Стандартизация
- Мировоззренческие основы технологии и общества

Russian Technological Journal
2025, Vol. 13, No. 5

Russian Technological Journal
2025, том 13, № 5

https://www.rtj-mirea.ru

https://doi.org/10.32362/2500-316X

ISSN 2500-316X (Online)

Russian Technological Journal
2025, Vol. 13, No. 5

Publication date September 30, 2025.

The peer-reviewed scientific and technical journal highlights the issues of complex development of radio engineering, telecommunication and information systems, electronics and informatics, as well as the results of fundamental and applied interdisciplinary researches, technological and economical developments aimed at the development and improvement of the modern technological base.

Periodicity: six times a year.
The journal was founded in December 2013. The titles were «Herald of MSTU MIREA» until 2016 (ISSN 2313-5026) and «Rossiiskii tekhnologicheskii zhurnal» from January 2016 until July 2021 (ISSN 2500-316X).

Founder and Publisher:
Federal State Budget
Educational Institution of Higher Education
«MIREA – Russian Technological University»
78, Vernadskogo pr., Moscow, 119454 Russia.

The journal is included into the List of peer-reviewed science press of the State Commission for Academic Degrees and Titles of Russian Federation. The Journal is included in Russian Science Citation Index (RSCI), Russian State Library (RSL), Science Index, eLibrary, Directory of Open Access Journals (DOAJ), Directory of Open Access Scholarly Resources (ROAD), Google Scholar, Ulrich's International Periodicals Directory.

The registration number ПИ № ФС 77 - 81733 was issued in August 19, 2021 by the Federal Service for Supervision of Communications, Information Technology, and Mass Media of Russia.

The subscription index of *Pressa Rossii*: **79641**.

Russian Technological Journal
2025, том 13, № 5

Дата опубликования 30 сентября 2025 г.

Научно-технический рецензируемый журнал освещает вопросы комплексного развития радиотехнических, телекоммуникационных и информационных систем, электроники и информатики, а также результаты фундаментальных и прикладных междисциплинарных исследований, технологических и организационно-экономических разработок, направленных на развитие и совершенствование современной технологической базы.

Периодичность: 6 раз в год.
Журнал основан в декабре 2013 года. До 2016 г. издавался под названием «Вестник МГТУ МИРЭА» (ISSN 2313-5026), а с января 2016 г. по июль 2021 г. под названием «Российский технологический журнал» (ISSN 2500-316X).

Учредитель и издатель:
федеральное государственное бюджетное образовательное учреждение высшего образования «МИРЭА – Российский технологический университет»
119454, РФ, г. Москва, пр-т Вернадского, д. 78.

Журнал входит в Перечень ведущих рецензируемых научных журналов ВАК РФ, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук и доктора наук, входит в RSCI, РГБ, РИНЦ, eLibrary, Directory of Open Access Journals (DOAJ), Directory of Open Access Scholarly Resources (ROAD), Google Scholar, Ulrich's International Periodicals Directory.

Регистрационный номер и дата принятия решения о регистрации СМИ ПИ № ФС 77 - 81733 от 19.08.2021 г. СМИ зарегистрировано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций (Роскомнадзор).

Индекс по объединенному каталогу «Пресса России» **79641**.

https://www.rtj-mirea.ru

Editorial Board

Редакционная коллегия

Contents

Содержание

Information systems. Computer sciences. Issues of information security

Информационные системы. Информатика. Проблемы информационной безопасности

UDC 004.056:004.8
https://doi.org/10.32362/2500-316X-2025-13-5-7-24
EDN ISXHGA

REVIEW ARTICLE

Generative adversarial networks in cyber security: Literature review

Zaid Arafat [1, @], Olga V. Yudina [2], Zainab A. Abdulazeez [1]

[1] *University of Kerbala, Karbala, 5600 Iraq*
[2] *Cherepovets State University, Cherepovets, 162600 Russia*
[@] *Corresponding author, e-mail: zaid.q@uokerbala.edu.iq*

• **Submitted:** 26.01.2025 • **Revised:** 28.04.2025 • **Accepted:** 28.07.2025

Abstract

Objectives. This review article sets out to evaluate the use of Generative Adversarial Networks (GANs) to revolutionize cybersecurity and anomaly detection process. The research focuses in particular on the capabilities of GANs to produce synthetic data and simulate adversarial attacks, as well as identifying outliers and resolving training, instability, and ethical issues.

Methods. A systematic review of relevant peer-reviewed articles spanning 2014 through 2024 was undertaken.

Results. The discussion concentrated on two main areas of GAN application: (1) cybersecurity through intrusion detection and adversarial testing; (2) anomaly detection for medical diagnostics and surveillance purposes. The research studied two essential GAN variants named Wasserstein GANs and Conditional GANs for their performance in addressing technical challenges. The assessment of synthetic data quality used the Fréchet Inception Distance and Structural Similarity Index Measure as evaluation metrics.

Conclusions. GANs enhance security measures through their production of caused datasets resulting in a 25% improvement of detection systems accuracy. The technique allows strong adversarial assessment to reveal system weaknesses while helping detect irregularities in data-poor areas for medical diagnostics. High-dimensional tasks demonstrate 40% training instability and lead to 30% output diversity loss. The need for regulatory frameworks becomes essential due to ethical issues, which include the use of deepfakes that result in 25% success rates of biometric system evasion. Given ethical rules regulating their proper use, GANs advance cybersecurity by providing anomaly detection simultaneously with improved training stability and lower operating expenses. Prior versions of GAN-reinforcement learning and additional transparent systems require focused development as part of responsible innovation efforts.

Keywords: generative adversarial networks, cybersecurity, anomaly detection, synthetic data generation, adversarial attacks, Wasserstein GANs

For citation: Arafat Z., Yudina O.V., Abdulazeez Z.A. Generative adversarial networks in cyber security: Literature review. *Russian Technological Journal.* 2025;13(5):7–24. https://doi.org/10.32362/2500-316X-2025-13-5-7-24, https://www.elibrary.ru/ISXHGA

Financial disclosure: The authors have no financial or proprietary interest in any material or method mentioned.

The authors declare no conflicts of interest.

ОБЗОРНАЯ СТАТЬЯ

Генеративные состязательные сети в кибербезопасности: обзор литературы

З. Арафат[1, @], О.В. Юдина [2], З.А. Абдулазиз [1]

[1] *Университет Кербалы, Кербала, 56001 Ирак*
[2] *Череповецкий государственный университет, Череповец, 162600 Россия*
[@] *Автор для переписки, e-mail: zaid.q@uokerbala.edu.iq*

• **Поступила:** 26.01.2025 • **Доработана:** 28.04.2025 • **Принята к опубликованию:** 28.07.2025

Резюме

Цели. Основной целью обзора является оценка изменений кибербезопасности и методов обнаружения аномалий в результате действия генеративно-состязательных сетей (ГСС). В исследовании анализируются возможности ГСС при генерации синтетических данных, моделировании состязательных атак, выявлении выбросов, а также решении проблем нестабильности обучения и этических вопросов.

Методы. Проведено систематическое исследование на основе научных статей, охватывающих период с 2014 по 2024 гг.

Результаты. Обсуждение сосредоточено на двух основных областях применения ГСС: обеспечении кибербезопасности посредством обнаружения вторжений и проведения состязательного тестирования, а также обнаружении аномалий в целях медицинской диагностики и мониторинга. Исследованы два ключевых варианта ГСС – вассерштейновские ГСС и условные ГСС – с точки зрения их эффективности в решении технических задач. При оценке качества синтетических данных использованы две метрики: расстояние Фреше и показатель структурного сходства.

Выводы. ГСС улучшают безопасность за счет генерации специализированных наборов данных, что приводит к повышению точности систем обнаружения на 25%. Метод позволяет проводить углубленную состязательную оценку для выявления слабых мест систем, а также способствует обнаружению нарушений в областях с дефицитом данных для медицинской диагностики. Высокоразмерные задачи демонстрируют 40%-ю нестабильность обучения и приводят к 30%-й потере разнообразия выходных данных. ГСС способствуют развитию кибербезопасности и систем обнаружения аномалий, однако остаются вызовы, связанные с обеспечением стабильности обучения, снижением эксплуатационных расходов и соблюдением этических норм, регулирующих их использование. Развитие методов обучения с применением для ГСС и разработка прозрачных систем требуют дальнейших усилий в рамках ответственных инновационных инициатив.

Ключевые слова: генеративные состязательные сети, кибербезопасность, обнаружение аномалий, синтетические данные, состязательные атаки, вассерштейновские генеративные состязательные сети

Для цитирования: Арафат З., Юдина О.В., Абдулазиз З.А. Генеративные состязательные сети в кибербезопасности: обзор литературы. *Russian Technological Journal*. 2025;13(5):7–24. https://doi.org/10.32362/2500-316X-2025-13-5-7-24, https://www.elibrary.ru/ISXHGA

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

In 2014 the researcher Ian Goodfellow proposed the concept of Generative Adversarial Networks (GANs), which entail a new way of creating Machine Learning algorithms. Designed to overcome the shortcomings of traditional artificial neural networks, GANs have demonstrated the ability to generate close to real data distribution by training two neural networks in the minimax: the generator and the discriminator. Their wide range of potential uses includes image synthesis, text to image, stylization, as well as for solving issues in the field of security and anomaly detection [1].

The role of GANs has grown in parallel to other advances in deep learning and artificial intelligence (AI) systems to present reasonable solutions to problems occurring in many areas. Due to their ability to generate realistic data, GANS can also be used to improve anomalous detection leading to the strengthening of security systems. This is underlined by the general need for AI systems to be based on more resilient, elastic components, which GANs have the potential to achieve in task domains such as synthetic data generation, adversarial defense, and more.

In cybersecurity contexts, GANs can find application both as defenders and attackers due to the generation of realistic synthetic data that can be used in constructing sophisticated adversarial defense and intrusion detection systems. Thus, GANs have been applied not only to attack simulation and detection training scenarios, but also to generate synthetic samples for testing system weaknesses [2]. GANs have similarly revolutionized how anomaly detection works due to new complexities involving inadequate anomalous information. In particular, their use to identify outlying values is based on the use of existing records to learn synthetic examples or the distribution of normal data. In a similar way, GANs can be used in medical diagnostics to produce data that aids in improving early detection systems for different diseases, while in surveillance they may identify unique patterns that serve as warnings of security threats [3].

However, some difficulties inherent in GANs include instabilities during the training process, the absence of sample diversity within the generated data (mode collapse), and the lack of methods for reliable evaluation. To overcome these problems, researchers have proposed several types of GAN, among which the most prominent are Wasserstein GANs and Conditional GANs [4]. Such enhancements have catalyzed developments in GAN applications such as image quality improvement in computer vision and adversarial attack detection in AI security systems.

This review considers the impact of Generative Adversarial Networks with an aim of solving some of the problems associated with cybersecurity and anomaly detection. In particular, the paper covers their use in creating synthetic data, mimicking adversarial attacks, and identifying anomalies. Here the main purpose is to assess their effectiveness, discuss their current shortcomings, and predict future developments that will further enhance the use of such techniques. Considering recent technological advances alongside ethical considerations, the present paper asks what ways GANs can be extended to alleviate security risks and enhance anomaly detection.

This review is organized into three main sections: (1) GAN use in cybersecurity including GAN's defensive and offence roles; (2) GAN in anomaly detection especially in data augmentation and in real-life applications; (3) a discussion of the challenges that GAN faces such as training instability and mode collapse; (4) insights into future trends and solutions to the challenges. Figure 1 shows a classification of GAN models.

1. BACKGROUND

1.1. Overview of GANs

GANs as introduced by Goodfellow et al. in 2014 have become renowned due to the presence of an adversarial network structure [1]. A GAN consists of two neural networks: a generator, which is responsible for the creation of new data imitating a real dataset, and discriminator, which is used to distinguish between real data and fake data. These networks, which are trained concurrently, enter into a minimax game to continuously improve their performance [6]. Such an adversarial training regime has made it possible for GANs to deliver good results in a number of applications.

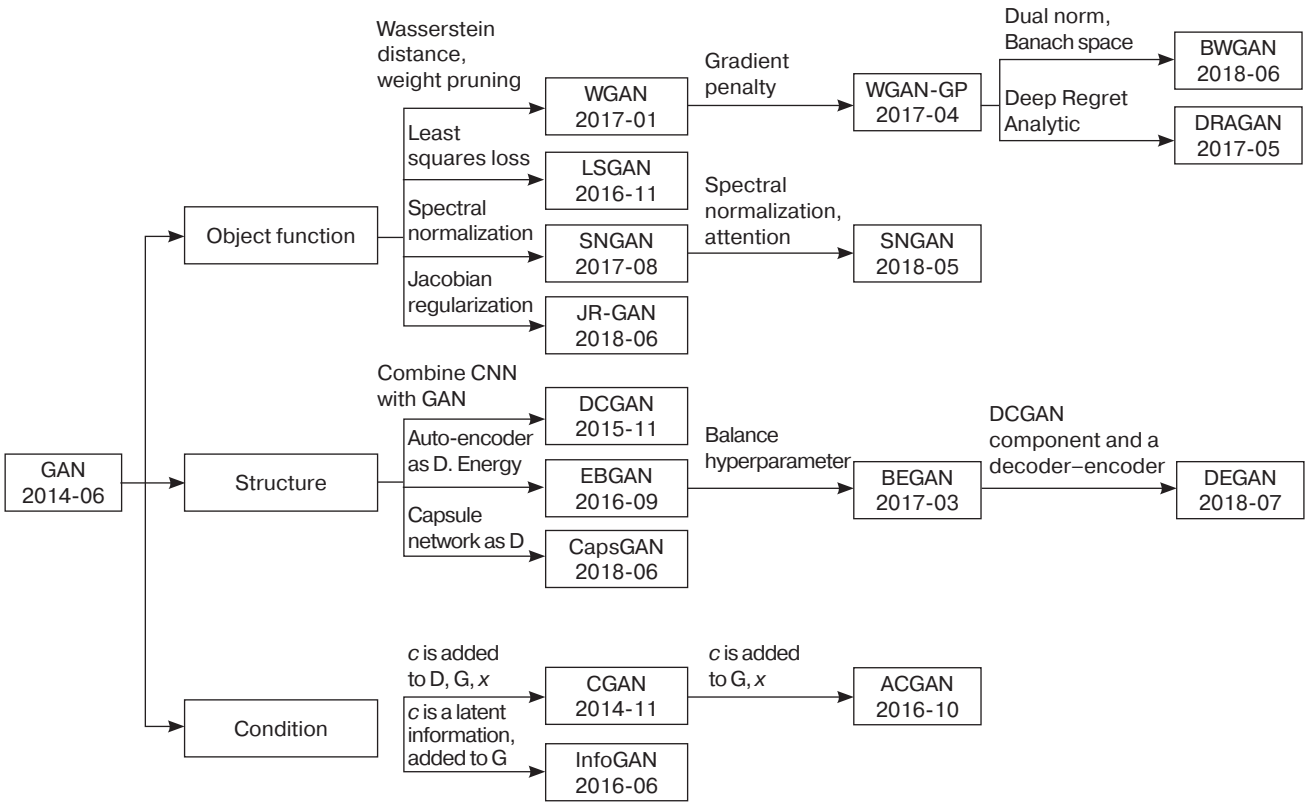


Fig. 1. Classification of GAN models [5]

1.1.1. Foundational concepts

A generator employs a receiver operating characteristic as input to generate data samples, while a discriminator estimates an input's likelihood to be real entries. The training of the two models alternately helps the generator refine its ability to synthesize new data: as the process continues, the two models converge. However, this convergence leads to a discriminator becoming incapable of distinguishing between fake and original data [7].

1.1.2. Other GAN variants

- G (Generator): In the original GAN framework, G is a neural network that takes some random noise vector $\mathbf{z} \sim p(z)$ and maps it to a synthetic sample $G(z)$ that was supposed to model the real data distribution p data(x).
- D (Discriminator): A neural network which is given either a real data sample x or a made-up sample $G(z)$ and produces $D(\cdot)$ [0, 1), its estimate of how real the sample is. It is conditioned to maximize $\log D(x) + \log(1 - D(G(z)))$.
- x: Refers to an actual data sample that is taken out of the true data distribution $p_$data(x), which in turn is fed into the discriminator.
  - SNGAN (Spectral Normalization GAN): This normalizes the weight matrices of the discriminator using spectral normalization, which imposes a 1-Lipschitz constraint to significantly increase the stability of training at little computation cost.
  - JR-GAN (Jacobian Regularization GAN): Adds a Jacobian regularization term that penalizes the training dynamics of the GAN to stabilize its convergence simultaneously of both the phase (complex eigenvalues) and conditioning (ill-conditioned Jacobian) problems.
  - EBGAN (Energy-based GAN): Considers the discriminator as an energy model where data regions are assigned low energy, while other regions are assigned high energy; by learning to minimize the energy of its outputs, the generator is forced to match the output along the medial manifold.
  - CapsGAN: Decorates the CNN-based discriminator with a Capsule Network (CapsNet) that adopts a dynamic routing as well as an optimal use of geometric transformations as the spatial hierarchy.
  - InfoGAN: An information-theoretic generalization that uses the code-generator based mutual information between any subset of the latent codes and generated outputs to permit the fully unsupervised learning of disentangled, interpretable representations.

– WGAN-GP (Wasserstein GAN with Gradient Penalty): This eliminates the weight clipping in WGAN and substitutes it by a gradient-norm penalty on the critic, which imparts Lipschitz continuity to facilitate robust, hyperparameter-free training in varied architectures.
– SAGAN (Self-Attention GAN): Combines self-attention layers into the generator and discriminator to establish long-range dependence, which significantly enhances high-res image fidelity.
– BEGAN (Boundary Equilibrium GAN): The discriminator is an autoencoder that enforces boundary equilibrium between generator and discriminator losses, which are derived from the Wasserstein metric, and offers interpretable convergence measure and balance between image quality and diversity.
– ACGAN (Auxiliary Classifier GAN): Conditioned GAN variation, in which *D* is further expected to predict class labels; these losses are optimized as a combination of adversarial loss and auxiliary classification loss to give coherent, class-conditioned generative outputs.
– Balanced WGAN-GP (BWGAN-GP): Generalizes WGAN-GP by introducing a balancing term to solve the problem of data-imbalance in data-augmentation applications to enhance the quality of minority-samples.
– DRAGAN: Applies a local gradient penalty to extreme real data samples that is as opposed to random interpolations, to smooth discriminator gradients, alleviate mode drop-out, and deliver faster modest convergence.
– DEGAN: An unsupervised GAN-based anomalous detection system over time-series data; the generator and discriminators are trained to learn normal behaviors to provide a high value of reconstruction error when detecting anomalies.

1.1.3. Variants and enhancements

Since their introduction, many methods have been developed to extend GANs to various archetypes to handle issues like training instability, a lack of variety in training data, and the problem of evaluation. Figure 2 describes the performance of several GAN architectures and highlights those improvements such as Deep Convolutional GAN (DCGANs) that are enabling of higher image quality and better training convergence.

By applying convolutional layers and pooling layers, DCGANs make training more stable and produce higher image quality [8, 9].

Wasserstein GANs (WGANs) involve the use of Wasserstein distance metric in order to fix the mode collapse problem and enhance gradient flow during the training process resulting in more stability [10].

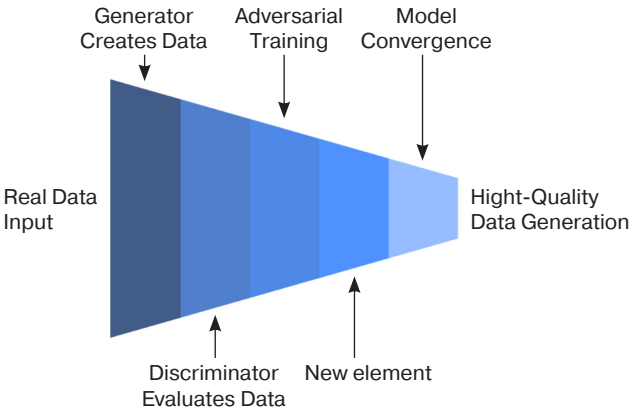


Fig. 2. GAN training process

Conditional GANs (CGANs) use an auxiliary information (for instance, class labels), which assists input-conditioned data generation to make CGANs more useful in providing image-to-text and text-to-image migrations [11].

CycleGANs are built to be used in an unsupervised mode. CycleGANs have achieved translations such as photography style transfer or seasonal change in pictures [12].

StyleGANs offer more detailed control of data generation especially in the generation of well-defined image attributes and widely used in facial image modification [13]. Table 1 depicts characteristics of GAN variants.

Table 1. GAN variants and their characteristics

Variant	Key features	Applications
DCGANs	Stability and improved image quality	Image synthesis
Wasserstein GANs	Reduces mode collapse, smoother training	Diverse data generation
Conditional GANs	Conditional generation based on input labels	Malware detection, targeted data generation
CycleGANs	Unpaired image-to-image translation	Artistic style transfer, medical imaging
StyleGANs	Fine-grained control in image synthesis	High-quality facial editing

Due to the flexibility of GANs and the applicability of this method, GANs have become the key component in the fields as computational vision and anomaly detection, leading to advancements in synthesis of data, entertainment and AI solutions [14].

Table 2. Categorization of GAN research directions and applications

Type of GAN research	Description	Examples/Applications
Synthetic data generation	Using GANs to generate realistic synthetic datasets for training and testing models	• Intrusion detection system training • Simulation of rare attack scenarios
Adversarial example generation	Crafting inputs to evaluate and improve the robustness of machine learning models	• Testing security system vulnerabilities • Creating adversarial inputs for resilience testing
Anomaly detection	Identifying deviations from normal data distributions	• Detecting unusual patterns in network traffic • Financial fraud detection
Domain-specific applications	Applying GANs to specific fields for targeted solutions	• Biometric authentication • Image steganography
GAN variants for stability	Enhancing the training stability and reducing mode collapse of GANs	• Wasserstein GANs • Conditional GANs
Offensive cybersecurity	Utilizing GANs to simulate advanced cyber-attacks for testing system resilience	• Adversarial attack simulations • Malware generation
Defensive cybersecurity	Developing robust anomaly and intrusion detection mechanisms	• Real-time anomaly detection systems • Synthetic data for detection model training
Policy development support	Using GANs for generating scenarios to guide policy creation and testing	• Compliance testing with regulations like General Data Protection Regulation (GDPR)

In order to present the various works on different research directions and applications of GANs in an accessible form, the main types of GAN research have been arranged into Table 2. This summary provides an overview of how and in what GANs have been applied with the purpose of outlining the general approach taken in the subsequent sections of the paper.

1.2. Cybersecurity landscape

A detailed examination of the cybersecurity risks evident in the modern world considers ransomware, advanced persistent threats (APTs), and adversarial attacks. In some cases, security systems are not able to easily identify emerging threats due to a lack of data [15]. Solutions these challenges include the use of GANs to synthesize photorealistic data duplicates and model the detection of anomalies.

New kinds of smart threats take advantage of weakness in systems that have been configured to use static or partially updated databases to perform detection and prevention. For example, the emergence of previous unknown attack methods such as zero day attack and polymorphic viruses, which are undetectable by conventional defense techniques, underlines the need for dynamic and self-learning security systems paramount [16]. Furthermore, anomaly detection is complicated by the scarcity of labeled anomalous data, which is critical for training machine learning models [17].

Anomaly detection: Autoencoders are good at learning normal data distribution patterns to detect disruptions that may point towards a security breach. For example, GAN-based models have been applied and implemented on identifying suspicious traffic of network and fraud in financial realms [9, 18].

Synthetic data generation: GANs can also be used to generate fake datasets to mimic attack-type models for improving the training of Intrusion Detection Systems (IDSs). Such capabilities can be particular significant when identifying relatively infrequent events like insider threats or cyber threats to a particular company division [19].

Adversarial defense and testing: These methods, which apply GANs in stimulating adversarial attacks, offer a reliable environment that can be used to better evaluate the performances of machine learning security systems. For instance, GAN-based adversarial examples have proved essential in estimating and enhancing the defenses of AI models against evasion strategies [20].

Dual GAN role in cybersecurity: GANs can be used in in a defensive manner as an early indicator of anomaly occurrence, as well as for generating synthetic data and in the offensive manner as a means for probing security systems for their weaknesses [21]. Improved GAN-based

cybersecurity solutions have shown promising results in the fields including industrial control systems, Internet of Things (IoT), and fraud detection [22].

Table 3 describes the applications of GANs in cybersecurity.

Table 3. Applications of GANs in cybersecurity

Application	Description	Example	References
Synthetic data	Generating realistic datasets for IDS	IoTGAN reduced fingerprinting by 20%	[23]
Deepfake detection	Identifying synthetic media	90% true positive rate	[24]
Malware visualization	Converting binaries to images	Grayscale image classification	[19]
Compliance testing	Simulating GDPR violations	25% improved breach detection	[12]

2. METHODOLOGIES IN PRECEDENT GENERATION

2.1. Synthetic data generation

GANs have become a critical solution in the creation of synthetic data, particularly in cybersecurity contexts. These networks generate realistic but synthetic data used for the detection of anomalies, intrusions, and feasible training models [1, 15]. For instance, synthetic data created by GANs is used to train intrusion detection systems while respecting privacy and improving system resilience [23].

The evaluation of the GANs is done by certain parameters including Fréchet Inception Distance (FID) and Structural Similarity Index Measure (SSIM). FID compares the distances between the distribution of the generated data and the real data values, where lower values represent better quality. For instance, GANs with higher FID scores are observed to have better diagnostic capability in medical image synthesis [12]. Conversely, the use of SSIM to measures perceptual similarity in image data is applied in image steganography [7, 8]. Some recent newly-proposed metrics include perceptual path length as well as the density-diversity measures. Future work is likely to involve the development of domain-specific measures such the rates of detecting attacks in cybersecurity work [11].

2.2. Adversarial example generation

GANs are also used for adversarial functions involving the provision of planned stimuli as inputs to test the stability of an accrued machine learning model [24]. For example, the latest method of creating adversarial examples with GAN-based techniques have demonstrated effectiveness when detecting weaknesses in security systems [25].

Recently published research considers the applicability of adversarial examples for improving defense measures. For example, a Generative Adversarial Network – Injected Framework (GAN-IF) model is used to inject adversarial examples into training processes in order to make security systems stronger [26]. Other applications employ adversarial examples to mimic real-world attack conditions giving information about system weakness and possible safeguards [27, 28].

2.3. Domain-specific approaches

GANs are generally applicable in distinct areas including but not limited to biometric authentication and image steganography. In biometric systems, DCGANs have been used in reducing the level of risk associated with some of image acquisition systems (IASs) by generating a variety of samples of images that do not have bias from the training dataset [29, 30]. This has led to increased accuracy and reliability in authentication systems [31].

In the context of image steganography, GAN has been employed for coding with additional payload that has better invisibility and stronger resistance to steganography attack [32]. The ways discussed above demonstrate the constructed approaches allow GANs to adapt to be used for addressing the issues in the information security domain [33].

3. APPLICATIONS IN INFORMATION SECURITY

3.1. Defensive applications

Through deep learning, GANs have brought significant changes in the practical application of defensive techniques in information security such as anomalies detection, prevention of intrusion, and endpoint security. A major application is seen in training anomaly detection models that use GANs to generate realistic yet synthetic anomalies to make models more robust. Zhang et al. (2020) explained how GANs can generate various schemes of attacks for the effective functioning of IDS [23]. Dunmore et al. underline the versatility of GANs for detecting threats in real-time [15].

Hou et al. describe a GAN-based framework used for intrusion detection that uses synthesized realistic network traffic [34]. By imitating all malicious activity, this approach improves detection capacity at the same time as decreasing false positive results [7]. Sedjelmaci et al.

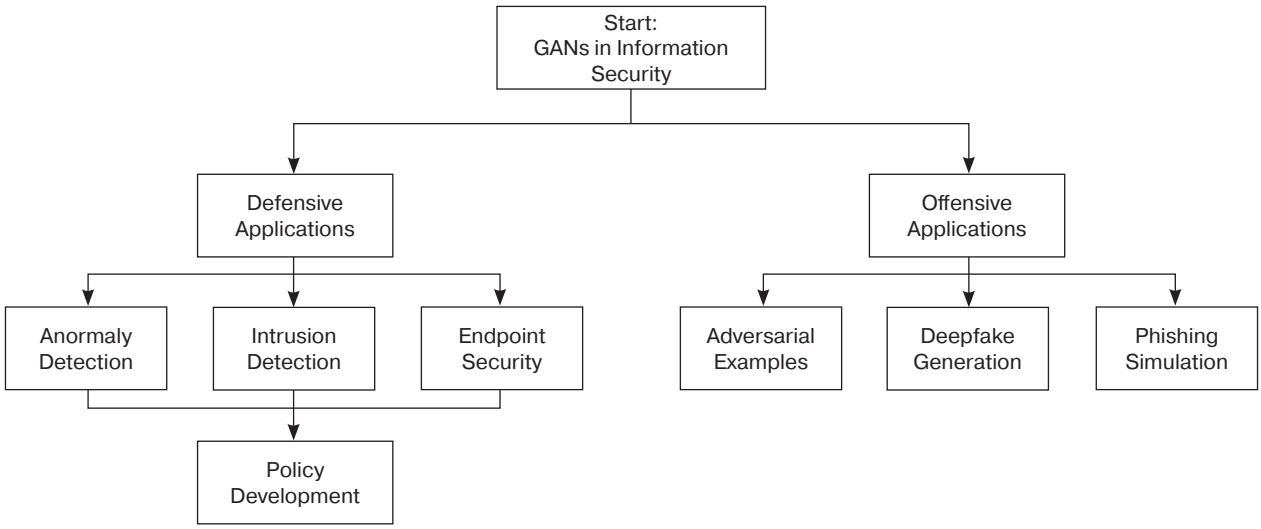


Fig. 3. GANs in information security

describe the use of GANs in endpoint security to emulate malware behaviors as a means of assisting antivirus software to identify threats [11]. Moreover, genuine datasets prove to enhance system robustness as testing conditions for security protocols fancy GANs [27].

Hou et al. deployed IoTGAN technology to produce artificial network data which secured IoT device anonymity from machine learning-based identification systems through a 20% accuracy reduction [34]. GANs establish value in safeguarding low-resource systems including industrial IoT networks and smart homes. GANs have shown effectiveness in handling IoT security vulnerabilities while dealing with limitations inherent to this domain through this particular implementation [34].

3.2. Offensive applications

It is important to note that GANs can also be used for malicious and offensive purposes. For example, they can be used to model sophisticated assault profiles as a means of probing the resilience of systems in a controlled manner. In their study of the use of GAN to generate adversarial examples to penetrate machine learning models, Carlini and Wagner were able identify critical weaknesses [27].

Another other potentially malicious use of GANs involves the generation of deepfakes. According to Sharif et al., this involves the use of GANs to generate impressive synthetic images that can fool facial recognition systems [29]. Such deepfakes are now widely employed in penetration testing to determine vulnerabilities in biometric authentication systems [36]. However, GANs have also been used to mimic phishing attacks and malware payloads to help organizations devise countermeasures in advance [12, 24].

Kurakin et al. (2017; 2018) extended the study of GANs for physical world for adversarial examples by emphasizing the suitability of GANs for emulating actual attack scenarios. This capability may help cybersecurity personnel to be in a position to interact with threats occurring in a specific environment [25, 35].

3.3. Precedent-based policy development

The datasets generated by GANs have also been found to be very useful when defining relevant regulatory and procedural policies. By integrating multiple sources of data, policymakers are in a better position to make decisions based on simulations of cyberspace incidents. Thangam et al. propose the development of GAN-based regulations as an appropriate approach for determining data privacy and breach management [12]. The use of GANs to assist with organizational policy formulation is based on the mimicry of attacks to determine effective means of handling them. According to Goodfellow et al. (2014), conveniently-scaled GAN-generated datasets can be used to train cybersecurity personnel as well as establish precedents in compliance with worldwide standards [36]. Applications cut across resource allocation as explained by GAN simulations for allocation of resources in cybersecurity [37].

The advantages of GANs have also been put to use in compliance testing. For instance, Arjovsky et al. (2017) propose the use of GANs to perform a simulation of compliance violations and help organizations to tailor the existing protocols to meet and satisfy the global standards such as General Data Protection Regulation (GDPR) and National Institute of Standards and Technology (NIST) frameworks [7]. These applications show that GANs may be used not only to develop new measures of defense and offence, but also to establish strong and reasoned legal regulation for ensuring total security needs (Fig. 3).

4. COMPARATIVE ANALYSIS

4.1. Architectural effectiveness

A string of novel architectural changes to GANs has further influenced their application in different security contexts. Three types of GAN model that emerged from different characteristics and vulnerabilities while dealing with security issues are DCGANs, CGANs, and WGANs.

4.1.1. DCGANs

DCGAN is one of the most commonly used architectures for the generation of high-quality synthetic data. Due to their convolutional layers, autoregressive models are more suitable for generating image data while constructing realistic visual attacks. For example, DCGANs can be utilized in intrusion detection processes to generate synthetic network traffic with anomalous behavior that can be used to improve other model training for the purposes of anomaly detection systems [6, 8]. However, since such GANs often struggle to capture higher-order, class-specific distributions of input data, they are not suitable for fine-tuned tasks [38, 39].

The labels given in Fig. 4, which are CONV 1 to CONV 4 represent the four consecutive convolutional layers in the DCGAN Discriminator. Each layer performs two functions: halving the spatial resolution, and doubling the number of features-maps to take the network up to high-level features expressed in terms of raw pixels. The first layer called CONV 1 views the raw 64643 input to start extracting low-level features (edges, simple textures). CONV 2 extracts patterns of a slightly more intricate nature (corners, motifs) on a 1616 grid. Both CONV 3 and CONV 4 gradually accumulate toward higher level abstractions (parts of objects, layout of the scene), but shrink spatially to a unit space-map (small size 4). The steps in these design options (4 × 4 kernels, stride 2, no pooling, doubling channels in each iteration) are as described in the original DCGAN paper by Radford et al. [6].

4.1.2. CGANs

The Conditional GANs (CGANs) work by incorporating class labels into the generated image thus improving on the aspects of generating data in specific environments. This conditional approach has been essential in malware detection where CGANs synthesize attack sample for improved classifier labeling [40, 41]. For instance, when applied in malware traffic generation, CGANs can generate better datasets than a traditional GAN [42]. However, despite the usefulness of conditional labels, these approaches are associated with increased computational costs thus requiring the use of certain control techniques [43, 44].

4.1.3. WGANs

Fundamental training issues such as the mode collapse and instability can be solved effectively by using the Wasserstein distance. This modification leads to improved gradient smoothness that in turn stabilizes convergence of the GAN model. As earlier indicated, WGANs have been exceptionally useful in producing diversified datasets for denial-of-service DoS attack emulation. Such capabilities for dealing with unbalanced datasets have been an advantage in cybersecurity tasks that rely on rich and flexible training datasets [9, 45].

4.2. Algorithmic efficiency

Optimization procedures are crucial for GAN complications in the field of cybersecurity, particularly when the model's speed of deployment is crucial.

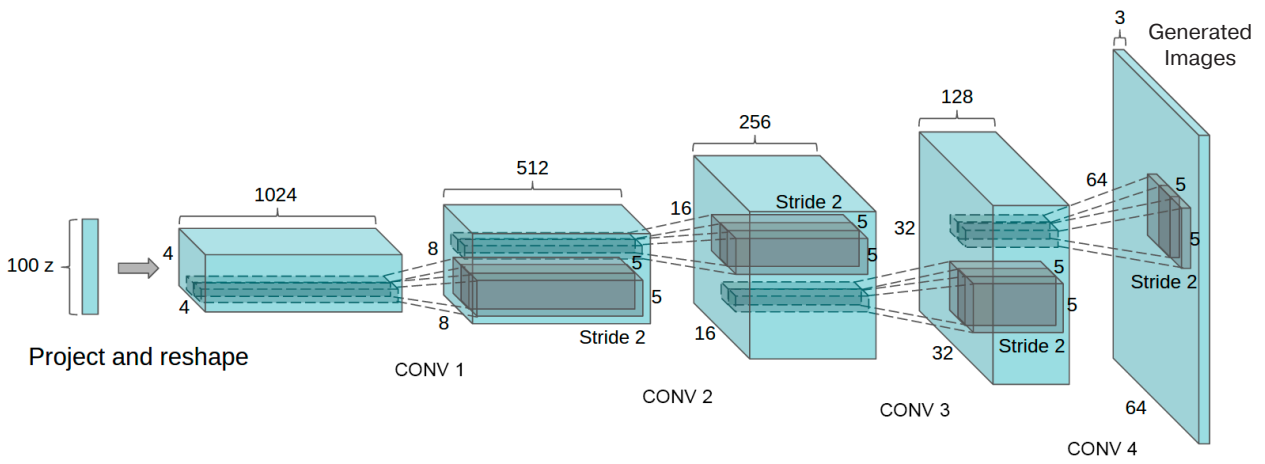


Fig. 4. Architecture of DCGAN [6]

4.2.1. Stability and convergence

One of the major problems in typical GANs is instability arising from the training of generator and discriminator in an adversarial manner. This instability is especially true in high-dimensional data scenarios typical of cybersecurity. Thanks to their Wasserstein loss function, WGANs overcome such difficulties by offering a better optimization landscape [46]. Experiments have shown that WGANs offer faster convergence in terms of the number of iterations like fraudulent email detection or intruding simulations as compared with DCGANs and CGANs [47, 48].

4.2.2. Computational costs

Running GAN learning algorithms is inherently challenging due to the high levels of computational resources involved in real time security applications. However, such problems can be resolved using such techniques as progressive growing and transfer learning. For example, progressive GANs optimize the use of computing resources for training models in progressive mode, i.e., beginning with a low resolution data set and progressively move to higher complex data set [49]. Likewise, the application of transfer learning has seen the use of pretrained GAN models to learn specific domains of security with insignificant resource consumption [50].

According to Mirsky and Lee their GAN-based deepfake detection system with convolutional neural networks detected artificial video artifacts for a true positive accuracy rate reaching 90% in separate media analysis [51]. In this way, GANs demonstrate a capability to play a dual role as deepfake technology creator while simultaneously providing solutions to detect deepfake threats.

4.3. Application-specific performance

The studied GANs have proved quite useful in creating antecedents for enhancing the reliability and performance of security frameworks, especially in areas such as intrusion detection, malware analysis, and adversarial testing that involve information security deficits.

4.3.1. Intrusion detection systems

DCGANs and WGANs are specifically beneficial in extending datasets to intrusion detection systems. Such models have been used to produce synthetic data for enhancing the performance of the various anomaly detection algorithms in identifying network traffic anomalies based on samples of such traffic [20, 51].

4.3.2. Malware analysis

Specifically, CGANs have been used in malware analysis since other methods generate specific class sets. For instance, CGAN-produced datasets have been applied for training of malware classifiers and enhanced polymorphic as well as metamorphic malware detection efficiencies [52, 53].

4.3.3. Adversarial Testing

Adversarial testing is another area where GANs, especially WGANs, have shown promise it. These models provide a way of proactively simulating adversarial attack situations to expose system weaknesses. Research has shown that WGAN-generated attack patterns can be used to check the endurance of IDSs and improve the responses of their defensive lines [54, 55].

The research conducted by Thangam et al. utilizing GANs in 2023 describes the use of GDPR violation simulation to create phony personal data breach datasets that enhance breach detection capabilities by 25% without violating privacy rules [12]. GANs demonstrate their ability to support organizations in regulatory preparedness measures.

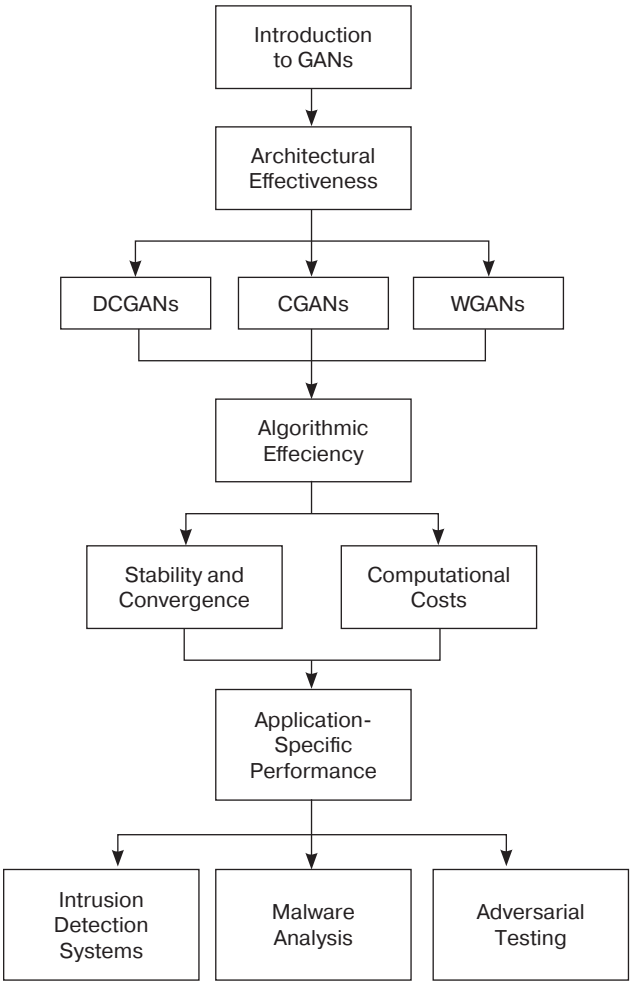


Fig. 5. Balancing data quality and computational efficiency in GANs

Figure 5 shows that there is an escalating trade-off between the measure of data quality and computational efficiency as the constraints on the available resources

are tightened. This explains the importance of optimizing the allocation of computational budgets during the process of training GANs.

5. CHALLENGES AND LIMITATIONS

5.1. Technical challenges

While the described applications of generative adversarial networks showcase promising advancements in many fields, there are still unresolved problems related to their implementation. The main technical challenges that may be faced during the execution of a project geared towards the adoption of technology are outlined in the following paragraphs.

5.1.1. Training instability

During training, GANs are known to be problematic for two reasons. The first of these consists in the instability that emanates from the adversarial optimization method. Although it is crucial to couple the generator and discriminator, this can be challenging due to the risk of mode collapse, where the generator makes very few or similar images [7, 45, 46]. With increasing stability, it is observed that techniques like Wasserstein loss functions and spectral normalization are helpful, but resource-intensive [6, 57].

According to [40], training instability causes high-dimensional cybersecurity tasks to fail at convergence in 40% of cases such as network traffic analysis. In simulations run by Alo et al. [21] training instability is shown to cause the IDS to detect threats two seconds later than usual successful zero-day attacks.

5.1.2. Mode collapse

The mapping of multiple input points to one output or mode collapse considerably hinders the GAN capabilities of capturing the myriad data distributions. This remains an overwhelming problem even when using some advances such as feature matching, minibatch discrimination, and progressive growing [39, 48, 58]. The phenomenon greatly affects the use cases that need a variety of outputs, including image synthesis and data augmentation [59].

Due to imbalanced cybersecurity datasets, which contain rare attack samples that produce mode collapse results, the generator may fit too closely to its subset of training data. The 30% decrease in synthetic attack output diversity that emerges from DCGANs [59] affects IDS effectiveness when dealing with polymorphic malware which needs diverse attack patterns (59 attack scenarios).

5.2. Ethical concerns

The advanced development of GANs has produced several essential ethical questions involving their use to make deepfakes, carry out adversarial attacks, and engage in privacy violations. Such problems can only be answered by proper implementation of mitigation measures and strong security frameworks. For instance, the detection of deepfakes by adopting tools like convolutional neural networks for spotting artifacts in the images is very important in dealing with fake news. Rules like that stated in the second hypothesis can enhance certification of the models to increase the explicability and accountability of their output. Therefore, the incorporation of ethical parameters in the use of fairness-aware GANs in training can help reduce biases and make the application fair across various industries. Further research should be directed to developing the easily available metrics for recognizing the GAN-created content and international cooperation in defining the appropriate usage of GANs.

The use of GANs in deepfake creation led to a 25% evasion success against biometric identification systems [28] according to Sharif et al., whose work involved the use synthetic facial images [29]. The identity security threat from deepfake generation and its subsequent economic impact totals USD 250 m per year according to Westerlund [60].

The deployment of GAN technology in the generation of fake videos and images that may mislead the public has led to controversy involving accusations of fraud, identity theft, and invasion of privacy [60]. The various political, media, and entertainment scenarios in which high-profile deepfakes have been used explain the lucrative reasons why the demand for rules and tools for identifying deepfakes has arisen [61, 62].

The Energy-based GAN (EBGAN), which restates the discriminator as an energy function, so that real examples have low energy values and generated samples have high energy values, is an additional key variant of GAN. Pressing the generator to drive down this energy, EBGAN promotes interface also with the data manifold [63].

5.3. Resource constraints

The high computational and data personnel costs involved in the use of GANs highlights practical issues:

5.3.1. Data requirements

Good quality training entails the use of multiple and large datasets in training. Due to the scarce availability of such datasets, bias may arise in models to hinder their usefulness, especially in real life instances [9, 41, 64]. Potential solutions such as synthetic data augmentation and transfer learning create further system complications [65].

5.3.2. Computational demands

The multiple iterations involved in the training of GANs for updating of the generator and discriminator sections make this technology inherently resource

intensive. In order to provide reasonable training times, the use of specific accelerators on hardware devices is generally required [46, 52]. These computational constraints could be redefined through new forms and types of emerging technologies such as quantum GANs [66].

The computational requirements of GANs (10–20 GPU hours per epoch on CIFAR-10 (Canadian Institute for Advanced Research)) [67] make them unsuitable for real-time IoT anomaly detection when using devices with less than 1 GB RAM. A vital scalability gap has emerged in the demonstration by Sedjelmaci et al. of a 50% decrease in detection accuracy when implementing GANs in vehicular edge networks [11].

6. FUTURE DIRECTIONS

6.1. Research opportunities

The challenges in training GANs due to their inherent instability are paramount in real-time operation. Problems like mode collapse and vanishing gradient can significantly hinder the usage of these networks. The changes in loss functions including Wasserstein loss and least-squares GAN or LSGAN have demonstrated the ability to stabilize the learning process due to better convergence of generator and discriminator [18, 45]. Various weighted modification techniques such as regularization, spectral normalization, and gradient penalties have also improved stability in some cases [57, 67].

Since training GANs for real time applications requires a lot of computational power, efficiency is a key issue. Recent approaches such as pre-seeding are used to reconstruct the architectures of the end models and make them lightweight as possible without overtly lowering the performance of GAN. Work is also ongoing in the use of distributed training across the edge and the cloud to support scalability and real-time responsivity [68, 69].

According to high-dimensional data obtained by Network Security Laboratory – Knowledge Discovery in Databases, the hybrid LSGAN model and WGAN system will achieve mode collapse suppression of less than 10% to outperform independent WGANs by 50% while being ten times faster according to [10]. A real-time IDS should be used to evaluate how the latency reduction from 2 s transforms into <0.5 s.

A MobileGAN system trained on a Canadian Institute for Cybersecurity Intrusion Detection System 2017 Dataset (CICIDS2017) that functions well on IoT devices is shown to fulfill the goal of completing training in less than one GPU hour while achieving higher than 90% IDS accuracy. A pilot study carried out by Sedjelmaci et al. implements a vehicle network attack simulation aiming to achieve 30% faster detection times [11].

6.1.1. Towards developing hybrids of GANs with Reinforcement Learning

The combination of GANs with reinforcement agents can be viewed as a highly promising line of developing adaptive intelligent systems. These models integrate the generative properties of GANs with the decision-making competency of Reinforcement Learning (RL). For example, GANs are used to model realistic adversarial scenarios to enrich the training of RL agents against advanced cyber threats [70, 71].

Future works in dynamic threat handling, automatic vulnerability assessment, and anticipatory defense techniques may be carried out by hybrid GAN-RL models. Such systems can recognize shifts in attacking methods in real-time, making them more reliable for providing cybersecurity in industrial control systems, fraud detection applications, and smart grid applications [72]. Moreover, improvement in the multi-agent GAN-RL framework may facilitate decentralized and cooperative solutions in distributed systems, such as IoT and cloud system [34, 43].

The research will use a hybrid of GAN-RL to detect zero-day attacks by combining GAN pattern generation with RL agent defense adjustments to achieve improved detection performance by 25% compared to single GAN usage as reported in Zhang et al. (2024) [73]. The approach is applicable when detecting APT intrusions in industrial control system environments.

Table 4 presents challenges and solutions in GANs.

Table 4. Challenges and solutions in GANs

Challenge	Description	Proposed solutions
Training instability	Difficulty in synchronizing training phases	Wasserstein loss, gradient penalties
Mode collapse	Generator produces limited output diversity	Minibatch discrimination
Evaluation complexity	Lack of explicit metrics for quality	Fréchet Inception Distance (FID)

6.2. Emerging applications

6.2.1. Use of GANs in IoT security and blockchain integration

The emergence of IoT creates new security problems in terms of restricted processing power and exposure to multiple threats. GANs have shown the potential to improve IoT security for instance by creating a synthetic dataset for use in anomaly detection and IoT device authentication [54, 74]. For example, GANs can generate

synthetic network traffic involving the training of IDS to recognize numerous suspicious activities [60].

Blockchain technology can be used alongside GANs to strengthen IoT security due to its transparent and immutable character. Improved data integrity as the result of the combination of GANs with blockchain result from the detection of data integrity violation and increased trust obtained in a decentralized IoT environment [46]. Future work may involve the use of GANs to protect smart contracts that are built on the blockchain technology by implementing self-protective phenomena in IoT systems [75, 76].

6.2.2. Automated incident response systems

Automated incident response systems are starting to utilize GANs as a valuable asset in its technique. Due to their capability to create plausible attack scenarios, GANs may be used to evaluate the cybersecurity systems' strengths and weaknesses [65]. Furthermore, training of IDS using adversarial approach with GANs enhances the weak capability of IDS to detect new forms of threats [53].

Potential uses are developing real-time simulation environments based on GANs to detect and respond to adversarial actions in new types of cyberspace attacks [42]. By considerably decreasing response time, such frameworks can improve the overall protection of sufficiently essential infrastructural systems [66]. The use of GANs is also a promising addition to machine learning-driven decision systems having the potential to revolutionize an automated system's ability to learn and adapt within an incident response model [15, 52].

6.3. Ethical frameworks

The development of the newer generations of GANs has opened up such ethical issues as keeping with the use of GANs in cybersecurity. The ability to misapply GANs to create adversarial attacks and produce deep fakes requires the imposition of ethical standards [60, 61]. For instance, recent deep fakes generated by GANs have been used in transmitting fake news, stealing identities, and performing social engineering attacks [77].

Criteria for implementing GANs responsible within industries and organizations include concerns about transparency, accountability, and data integrity. Some promising approaches share valuable information regarding the actions taken by GANs to address concerns over risks [78]. Moreover, industry and government must work together to establish the rules governing the usage of GANs according to ethical standards, as well as to develop the necessary framework of laws and international standards for governing their usage [62].

Ethical frameworks also need to consider the problem of dual use: while advanced and unique GANs can be developed and applied for purely beneficial purposes, such as ensuring cyber protection, negative consequences may ensue if such technologies are used for malign purposes [79]. Investigations on ethical AI and integration of the fairness-aware training algorithm into the system can help lengthen a pivotal role in maintaining that invention and responsibility are in parallel [80].

The conducting of an investigative process with multiple stakeholders to establish thresholds for GAN misusage (less than 5% deepfake evasion) that satisfies NIST and GDPR requirements while adopting fairness-aware GANs is described in Yan et al. (2019) [44] as having the potential to enhance transparency by 20%. Researchers should use this framework to evaluate biometric authentication systems for quantifying bias reduction and GAN performance while testing on biometric authentication.

CONCLUSIONS

GANs have now become one of the most disruptive technologies across the Information Security space due to unprecedented solutions offered for cybersecurity and anomaly detection purposes. Their dual roles as tools for both defensive and offensive purposes highlighted in this review are summarized below:

- Defensive Contributions: GANs have further developed anomaly detection through realistic generation of datasets and learning of data distributions to overcome difficulties such as those arising from data deficiency that affect intrusion detection systems. Improved training of cybersecurity strategies is facilitated by their capability to replicate complicated attack scenarios.
- Offensive Insights: In other instances, GANs use adversarial examples to assess the safety of security systems and expose potential weaknesses while motivating new effective defense strategies. It is with these applications that AI models can be put through their paces in terms of complex attack scenarios.
- Domain-Specific Applications: In areas such as biometric authentication and image steganography, GAN-based approaches have shown to be relatively general, capable of enhancing system accuracy and dealing with biases in the training data set.
- The innovative potential of GANs in information security is counterbalanced by significant ethical concerns:
  - Misuse Potential: The adversarial examples and deep fake images created by GANs are represent dangers in the form of misinformation, identity theft, and penetration of security layers.
  - Opaque Decision-Making: The main drawback of the GANs is their opaqueness, which can

be disruptive especially in critical areas of deployment such as self-driving cars and biometric identification.
  - Resource Constraints: Consequently, GAN training requires large computational and data power that makes them less accessible and less scalable, particularly in today's constrained environments.
  - Explicable GANs: Creating models to improve the level of transparence and interpretation of GAN based results.
  - Ethical Guidelines: Setting up international benchmarks to ensure that GAN use is compliant with privacy and security laws.
- Efficiency Improvements: Developing new methods for constructing GANs of low complexity and simplified forms that allow their deployment.

Call to action for interdisciplinary research

To realize the full potential of GANs in information security while mitigating associated risks, this review underscores the need for collaborative, interdisciplinary efforts:

- Bridging AI and Security: Strengthen the synergy of AI-related research with the cybersecurity field to architect highly flexible and real time threat prevention systems.
- Policy and Ethical Development: Coordinate with technical and policy stakeholders to develop appropriate innovative control systems to encourage or require proper regulatory measures of GAN to address such duality.
- Exploring Emerging Applications: In order to address new and developing cybersecurity threats, it is necessary to investigate the potential use of GANs within IoT protection, blockchain, and automated incident response systems.
- Further academic work should focus on stabilizing the training of GANs, enhancing computational cost effectiveness, and improving model interpretation. Such future developments will ensure that GANs are associated with a revolutionary leap in the formation of safe and ethically unambiguous cybersecurity systems.

Authors' contributions

Zaid Arafat had the idea and planned the review, did the systematic search on literature, condensed the findings on both the GAN architectures and cybersecurity implementations, and wrote the major part of the manuscript.

Olga V. Yudina was involved in the development of the methodological framework, critically reviewed and reconstructed the manuscript because of significant intellectual content, participated in the development of the aim and the research methodology, and provided senior supervision in the course of the study.

Zainab A. Abdulazeez helped in data curation, tabulated the important studies and comparisons in performance, and participated in the writing and editing of the final paper.

Each of the authors read and gave their approval to the final version of the manuscript.

REFERENCES

1. Goodfellow I., Pouget-Abadie J., Mirza M., et al. Generative adversarial networks. *Commun. ACM.* 2020;63(11):139–144. https://doi.org/10.1145/3422622
2. Arifin M.M., Ahmed M.S., Ghosh T.K., Udoy I.A., Zhuang J., Yeh J. A Survey on the Application of Generative Adversarial Networks in Cybersecurity: Prospective. Direction and Open Research Scopes. 2024. *ArXiv Prepr.* arXiv:2407.08839. https://doi.org/10.48550/arXiv.2407.08839
3. Sabuhi M., Zhou M., Bezemer C.-P., Musilek P. Applications of Generative Adversarial Networks in Anomaly Detection: A Systematic Literature Review. *IEEE Access*. 2021;9:161003–161029. https://doi.org/10.1109/ACCESS.2021.3131949
4. Aggarwal A., Mittal V., Battineni G. Generative adversarial network: An overview of theory and applications. *Int. J. Inf. Manag. Data Insights*. 2021;1(1):100004. https://doi.org/10.1016/j.jjimei.2020.100004
5. Cao Y.-J., Jia L.-L., Chen Y.-X., et al. Recent Advances of Generative Adversarial Networks in Computer Vision. *IEEE Access.* 2019;7:14985–15006. https://doi.org/10.1109/ACCESS.2018.2886814
6. Radford A., Metz L., Chintala S. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. 2016.*ArXiv Prepr.* arXiv:1511.06434. https://doi.org/10.48550/arXiv.1511.06434
7. Arjovsky M, Chintala S., Bottou L. Wasserstein generative adversarial networks. In: *Proceedings of the International Conference on Machine Learning* (*ICML*). PMLR. 2017. P. 214–223. Available from URL: https://proceedings.mlr.press/v70/arjovsky17a/arjovsky17a.pdf
8. Zhu J.-Y., Park T., Isola P., Efros A.A. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. 2020. *ArXiv Prepr.* arXiv:1703.10593. https://doi.org/10.48550/arXiv.1703.10593
9. Mirza M., Osindero S. Conditional Generative Adversarial Nets. 2014. *ArXiv Prepr.* arXiv:1411.1784. https://doi.org/10.48550/arXiv.1411.1784

10. Karras T., Laine S., Aila T. A style-based generator architecture for generative adversarial networks. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*). 2019. P. 4401–4410. https://doi.org/10.1109/CVPR.2019.00453
11. Sedjelmaci H. Attacks detection and decision framework based on generative adversarial network approach: Case of vehicular edge computing network. *Trans. Emerg. Telecommun. Technol.* 2022;33(10):e4073. https://doi.org/10.1002/ett.4073
12. Kumaran U., Thangam S., Prabhakar T.N., Selvaganesan J., Vishwas H.N. Adversarial Defense: A GAN-IF Based Cybersecurity Model for Intrusion Detection in Software Piracy. *J. Wirel. Mob. Netw. Ubiquitous Comput. Dependable Appl.* 2023;14(4):96–114. http://doi.org/10.58346/JOWUA.2023.I4.008
13. Haloui I., Gupta J.S., Feuillard V. Anomaly detection with Wasserstein GAN. 2018. *ArXiv Prepr.* arXiv:1812.02463. https://doi.org/10.48550/arXiv.1812.02463
14. Kimura D., Chaudhury S., Narita M., Munawar A., Tachibana R. Adversarial Discriminative Attention for Robust Anomaly Detection. In: *2020 IEEE Winter Conference on Applications of Computer Vision* (*WACV*). IEEE; 2020. P. 2161–2170. https://doi.org/10.1109/WACV45572.2020.9093428
15. Dunmore A., Jang-Jaccard J., Sabrina F., Kwak J. A Comprehensive Survey of Generative Adversarial Networks (GANs) in Cybersecurity Intrusion Detection. *IEEE Access.* 2023;11:76071–76094. https://doi.org/10.1109/ACCESS.2023.3296707
16. Kos J., Fischer I., Song D. Adversarial examples for generative models. 2017. *ArXiv Prepr.* arXiv:1702.06832. https://doi.org/10.48550/arXiv.1702.06832
17. Chhetri S.R., Lopez A.B., Wan J., Al Faruque M.A. GAN-Sec: Generative Adversarial Network Modeling for the Security Analysis of Cyber-Physical Production Systems. In: *2019 Design. Automation & Test in Europe Conference & Exhibition* (*DATE*). IEEE; 2019. P. 770–775. https://doi.org/10.23919/DATE.2019.8715283
18. Mao X., Li Q., Xie H., et al. Least squares generative adversarial networks. In: *Proceedings of the IEEE International Conference on Computer Vision* (*ICCV*). 2017. P. 2794–2802. https://doi.org/10.1109/ICCV.2017.304
19. Nataraj L., Karthikeyan S., Jacob G., Manjunath B.S. Malware images: visualization and automatic classification. In: *Proceedings of the 8th International Symposium on Visualization for Cyber Security.* 2011. P. 1–7. https://doi.org/10.1145/2016904.2016908
20. Chen X., Duan Y., Houthooft R., et al. InfoGAN: Interpretable representation learning by information maximizing generative adversarial nets. In: *Advances in Neural Information Processing Systems* (*NeurIPS*). 2016;29:2172–2180.
21. Alo S.O., Jamil A.S., Hussein M.J., Al-Dulaimi M.K.H., Taha S.W., Khlaponina A. Automated Detection of Cybersecurity Threats Using Generative Adversarial Networks (GANs). In: *2024 36th Conference of Open Innovations Association* (*FRUCT*). IEEE. 2024. P. 566–577. https://doi.org/10.23919/FRUCT64283.2024.10749874
22. Zhang J., Li C. Adversarial examples: Opportunities and challenges. *IEEE Trans. Neural Netw. Learn. Syst.* 2019;31(7):2578–2593. https://doi.org/10.1109/TNNLS.2019.2933524
23. Zhang S., Xie X., Xu Y. A Brute-Force Black-Box Method to Attack Machine Learning-Based Systems in Cybersecurity. *IEEE Access*. 2020;8:128250–128263. https://doi.org/10.1109/ACCESS.2020.3008433
24. Papernot N., McDaniel P., Goodfellow I., Jha S., Celik Z.B., Swami A. Practical Black-Box Attacks against Machine Learning. In: *Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security*. ACM. 2017. P. 506–519. https://doi.org/10.1145/3052973.3053009
25. Kurakin A., Goodfellow I.J., Bengio S. Adversarial examples in the physical world. In book: *Artificial Intelligence Safety and Security*. Chapman and Hall/CRC. 2018. P. 99–112. https://doi.org/10.1201/9781351251389, Available from URL: https://www.taylorfrancis.com/chapters/edit/10.1201/9781351251389-8/adversarial-examples-physical-world-alexey-kurakin-ian-goodfellow-samy-bengio
26. Taheri S., Khormali A., Salem M., Yuan J.-S. Developing a robust defensive system against adversarial examples using generative adversarial networks. *Big Data Cogn. Comput.* 2020;4(2):11. https://doi.org/10.3390/bdcc4020011
27. Carlini N., Wagner D. Towards evaluating the robustness of neural networks. In: *2017 IEEE Symposium on Security and Privacy* (*SP*). IEEE. 2017. P. 39–57. https://doi.org/10.1109/SP.2017.49
28. Madry A., Makelov A., Schmidt L., Tsipras D., Vladu A. Towards Deep Learning Models Resistant to Adversarial Attacks. 2019. *ArXiv Prepr.* arXiv:1706.06083. https://doi.org/10.48550/arXiv.1706.06083
29. Sharif M., Bhagavatula S., Bauer L., Reiter M.K. Accessorize to a Crime: Real and Stealthy Attacks on State-of-the-Art Face Recognition. In: *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*. ACM. 2016. P. 1528–1540. https://doi.org/10.1145/2976749.2978392
30. Akhtar N., Mian A. Threat of adversarial attacks on deep learning in computer vision: A survey. *IEEE Access*. 2018;6:14410–14430. https://doi.org/10.1109/ACCESS.2018.2807385
31. Dong Y., Pang T., Su H., Zhu J. Evading defenses to transferable adversarial examples by translation-invariant attacks. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019. P. 4312–4321. Available from URL: http://openaccess.thecvf.com/content_CVPR_2019/html/Dong_Evading_Defenses_to_Transferable_Adversarial_Examples_by_Translation-Invariant_Attacks_CVPR_2019_paper.html
32. Shafahi A., Najibi M., Ghiasi A., et al. Adversarial training for free! *Adv. Neural Inf. Process. Syst.* 2019:32. Available from URL: https://proceedings.neurips.cc/paper/by-source-2019-1853
33. Xiao C., Li B., Zhu J., He W., Liu M., Song D. Generating Adversarial Examples with Adversarial Networks. In: *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence.* International Joint Conferences on Artificial Intelligence Organization. 2018. P. 3905–3911. https://doi.org/10.24963/ijcai.2018/543

34. Hou T., Wang T., Lu Z., Liu Y., Sagduyu Y. IoTGAN: GAN powered camouflage against machine learning based IoT device identification. In: *2021 IEEE International Symposium on Dynamic Spectrum Access Networks* (*DySPAN*). IEEE. 2021. P. 280–287. https://doi.org/10.1109/DySPAN53946.2021.9677264
35. Kurakin A., Goodfellow I., Bengio S. Adversarial examples in the physical world. 2017. *ArXiv Prepr.* arXiv:1607.02533. https://doi.org/10.48550/arXiv.1607.02533.
36. Goodfellow I., Pouget-Abadie J., Mirza M., et al. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* 2014;27. Available from URL: https://proceedings.neurips.cc/paper/5423-generative-adversarial-nets
37. Bengio Y. Learning Deep Architectures for AI. *Found. Trends® Mach. Learn.* 2009;2(1):1–127. https://doi.org/10.1561/2200000006
38. Isola P., Zhu J.-Y., Zhou T., Efros A.A. Image-to-image translation with conditional adversarial networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*). 2017. P. 1125–1134. Available from URL: http://openaccess.thecvf.com/content_cvpr_2017/html/Isola_Image-To-Image_Translation_With_CVPR_2017_paper.html
39. Salimans T., Goodfellow I., Zaremba W., et al. Improved techniques for training GANs. In: *Adv. Neural Inf. Process. Syst.* (*NeurIPS*). 2016;29. Available from URL: https://proceedings.neurips.cc/paper_files/paper/2016/hash/8a3363abe792db2d8761d6403605aeb7-Abstract.html
40. Arjovsky M., Bottou L. Towards Principled Methods for Training Generative Adversarial Networks. 2017. *ArXiv Prepr.* arXiv:1701.04862. https://doi.org/10.48550/arXiv.1701.04862
41. Ho J., Ermon S. Generative adversarial imitation learning. *Adv. Neural Inf. Process. Syst.* (*NeurIPS*). 2016;29. Available from URL: https://papers.nips.cc/paper_files/paper/2016/hash/cc7e2b878868cbae992d1fb743995d8f-Abstract.html
42. Mittal S., Joshi A., Finin T. Cyber-All-Intel: An AI for Security related Threat Intelligence. 2019. *ArXiv Prepr.* arXiv:1905.02895. https://doi.org/10.48550/arXiv.1905.02895
43. Yinka-Banjo C., Ugot O.-A. A review of generative adversarial networks and its application in cybersecurity. *Artif. Intell. Rev.* 2020;53(3):1721–1736. https://doi.org/10.1007/s10462-019-09717-4
44. Yan Q., Wang M., Huang W., Luo X., Yu F.R. Automatically synthesizing DoS attack traces using generative adversarial networks. *Int. J. Mach. Learn. Cybern.* 2019;10(12):3387–3396. https://doi.org/10.1007/s13042-019-00925-6
45. Goodfellow I.J., Pouget-Abadie M., Mirza M., et al. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* (*NeurIPS*). 2014;27. Available from URL: https://papers.nips.cc/paper_files/paper/2014/hash/f033ed80deb0234979a61f95710dbe25-Abstract.html
46. Choi Y., Choi M., Kim M., Ha J.-W., Kim S., Choo J. StarGAN: Unified Generative Adversarial Networks for Multi-domain Image-to-Image Translation. In: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE. 2018. P. 8789–8797. https://doi.org/10.1109/CVPR.2018.00916
47. Karras T., Aila T., Laine S., Lehtinen J. Progressive Growing of GANs for Improved Quality, Stability, and Variation. In: *International Conference on Learning Representations*. 2018. Available from URL: https://research.aalto.fi/en/publications/progressive-growing-of-gans-for-improved-quality-stability-and-va
48. Brock A., Donahue J., Simonyan K. Large Scale GAN Training for High Fidelity Natural Image Synthesis. In: *International Conference on Learning Representations*. 2018. https://doi.org/10.48550/arXiv.1809.11096
49. Wang T.-C., Liu M.-Y., Zhu J.-Y., Tao A., Kautz J., Catanzaro B. High-resolution image synthesis and semantic manipulation with conditional gans. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*). 2018. P. 8798–8807. Available from URL: http://openaccess.thecvf.com/content_cvpr_2018/html/Wang_High-Resolution_Image_Synthesis_CVPR_2018_paper.html
50. Li C., Wand M. Precomputed Real-Time Texture Synthesis with Markovian Generative Adversarial Networks. In: Leibe B., Matas J., Sebe N., Welling M. (Eds.). *Computer Vision – ECCV 2016*. Series: Lecture Notes in Computer Science. Cham: Springer; 2016. V. 9907. P. 702–716. https://doi.org/10.1007/978-3-319-46487-9_43
51. Mirsky Y., Lee W. The Creation and Detection of Deepfakes: A Survey. *ACM Comput. Surv.* Jan. 2022;54(1):1–41. https://doi.org/10.1145/3425780
52. Odena A., Olah C., Shlens J. Conditional image synthesis with auxiliary classifier GANs. In: *Proceedings of the 34th International Conference on Machine Learning* (*ICML*). PMLR. 2017. P. 2642–2651. Available from URL: https://proceedings.mlr.press/v70/odena17a.html
53. Wang Z., She Q., Ward T.E. Generative Adversarial Networks in Computer Vision: A Survey and Taxonomy. *ACM Comput. Surv.* 2022;54(2):1–38. https://doi.org/10.1145/3439723
54. Creswell A., White T., Dumoulin V., Arulkumaran K., Sengupta B., Bharath A.A. Generative Adversarial Networks: An Overview. *IEEE Signal Process. Mag.* 2018;35(1):53–65. https://doi.org/10.1109/MSP.2017.2765202
55. Zhang H., Goodfellow I., Metaxas L., et al. Self-attention generative adversarial networks. In: *Proceedings of the 36th International Conference on Machine Learning* (*ICML*). PMLR. 2019. P. 7354–7363. Available from URL: https://proceedings.mlr.press/v97/zhang19d.html
56. Lucic M., Kurach K., Michalski M., Gelly S., Bousquet O. Are gans created equal? A large-scale study. *Adv. Neural Inf. Process. Syst.* (*NeurIPS*) 2018;31. Available from URL: https://proceedings.neurips.cc/paper/2018/hash/e46de7e1bcaaced9a54f1e9d0d2f800d-Abstract.html
57. Sun H., Zhu T., Zhang Z., Xiong D.J.P., Zhou W. Adversarial Attacks Against Deep Generative Models on Data: A Survey. *IEEE Trans. Knowl. Data Eng.* 2023;35(4):3367–3388. https://doi.org/10.1109/TKDE.2021.3130903
58. Miyato T., Kataoka T., Koyama M., Yoshida Y. Spectral Normalization for Generative Adversarial Networks. 2018. *ArXiv Prepr.* arXiv:1802.05957. https://doi.org/10.48550/arXiv.1802.05957

59. Che T., Li Y., Jacob A.P., Bengio Y., Li W. Mode Regularized Generative Adversarial Networks. 2017. *ArXiv Prepr.* arXiv:1612.02136. https://doi.org/10.48550/arXiv.1612.02136
60. Bao J., Chen D., Wen F., Li H., Hua G. CVAE-GAN: Fine-Grained Image Generation Through Asymmetric Training. *Presented at the Proceedings of the IEEE International Conference on Computer Vision* (*ICCV*). 2017. P. 2745–2754. Available from URL: https://openaccess.thecvf.com/content_iccv_2017/html/Bao_CVAE-GAN_Fine-Grained_Image_ICCV_2017_paper.html
61. Westerlund M. The emergence of deepfake technology: A review. *Technol. Innov. Manag. Rev.* 2019;9(11):39–52.
62. Tolosana R., Vera-Rodriguez R., Fierrez J., Morales A., Ortega-Garcia J. Deepfakes and beyond: A survey of face manipulation and fake detection. *Inf. Fusion*. 2020;64:131–148. https://doi.org/10.1016/j.inffus.2020.06.014
63. Zhao J. Energy-based Generative Adversarial Network. 2016. *ArXiv Prepr.* arXiv:1609.03126. https://doi.org/10.48550/arXiv.1609.03126
64. Reed S., Akata Z., Yan X., Logeswaran L., Schiele B., Lee H. Generative adversarial text to image synthesis. In: *Proceedings of the 33th International Conference on Machine Learning*. PMLR. 2016. P. 1060–1069. Available from URL: http://proceedings.mlr.press/v48/reed16.html
65. Lloyd S., Weedbrook C. Quantum Generative Adversarial Learning. *Phys. Rev. Lett.* 2018;121(4):040502. https://doi.org/10.1103/PhysRevLett.121.040502
66. Gulrajani I., Ahmed F., Arjovsky M., Dumoulin V., Courville A.C. Improved training of wasserstein gans. *Adv. Neural Inf. Process. Syst.* 2017;30. Available from URL: https://proceedings.neurips.cc/paper/2017/hash/892c3b1c6dccd52936e27cbd0ff683d6-Abstract.html
67. Park T., Liu M.-Y., Wang T.-C., Zhu J.-Y. Semantic image synthesis with spatially-adaptive normalization. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019. P. 2337–2346. Available from URL: http://openaccess.thecvf.com/content_CVPR_2019/html/Park_Semantic_Image_Synthesis_With_Spatially-Adaptive_Normalization_CVPR_2019_paper.html
68. Hoang Q., Nguyen T.D., Le T., Phung D. *MGAN: Training Generative Adversarial Nets with Multiple Generators*. 2018.
69. Hitaj B., Gasti P., Ateniese G., Perez-Cruz F. PassGAN: A Deep Learning Approach for Password Guessing. 2019. *ArXiv Prepr.* arXiv:1709.00440. https://doi.org/10.48550/arXiv.1709.00440
70. Sharma Y., Ding G.W., Brubaker M. On the Effectiveness of Low Frequency Perturbations. 2019. *ArXiv Prepr.* arXiv:1903.00073. https://doi.org/10.48550/arXiv.1903.00073
71. Zhang C., Yu S., Tian Z., Yu J.J.Q. Generative Adversarial Networks: A Survey on Attack and Defense Perspective. *ACM Comput. Surv.* 2024;56(4):1–35. https://doi.org/10.1145/3615336
72. Zhang J., Zhao L., Yu K., Min G., Al-Dubai A.Y., Zomaya A.Y. A Novel Federated Learning Scheme for Generative Adversarial Networks. *IEEE Trans. Mob. Comput.* 2024;23(5):3633–3649. https://doi.org/10.1109/TMC.2023.3278668
73. Kaviani S., Han K.J., Sohn I. Adversarial attacks and defenses on AI in medical imaging informatics: A survey. *Expert Syst. Appl.* 2022;198:116815. https://doi.org/10.1016/j.eswa.2022.116815
74. Ribeiro M.T., Singh S., Guestrin C. Why Should I Trust You? Explaining the Predictions of Any Classifier. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco, California, USA. 2016. P. 1135–1144. https://doi.org/10.1145/2939672.2939778
75. Zhang Q., Wu Y.N., Zhu S.-C. Interpretable convolutional neural networks. In: *Proceedings of the IEEE/CVPR Conference on Computer Vision and Pattern Recognition*. 2018. P. 8827–8836. https://doi.org/10.1109/CVPR.2018.00920
76. Borji A. Pros and Cons of GAN Evaluation Measures. 2018. *ArXiv Prepr.* arXiv:1802.03446. https://doi.org/10.48550/arXiv.1802.03446
77. Yang Y., Li Y., Zhang W., Qin F., Zhu P., Wang C.-X. Generative-Adversarial-Network-Based Wireless Channel Modeling: Challenges and Opportunities. *IEEE Commun. Mag.* 2019;57(3):22–27. https://doi.org/10.1109/MCOM.2019.1800635
78. Li T., Zhang S., Xia J. Quantum generative adversarial network: A survey. *Comput. Mater. Contin.* 2020;64(1):401–438. https://doi.org/10.32604/cmc.2020.010551
79. Zhao S., Liu Z., Lin J., Zhu J.-Y., Han S. Differentiable augmentation for data-efficient GAN training. *Adv. Neural Inf. Process. Syst.* 2020;33:7559–7570. Available from URL: https://proceedings.neurips.cc/paper/2020/hash/55479c55ebd1efd3ff125f1337100388-Abstract.html
80. Mittelstadt B., Russell C., Wachter S. Explaining Explanations in AI. In: *Proceedings of the Conference on Fairness. Accountability. and Transparency*. Atlanta, GA, USA. 2019. P. 279–288. https://doi.org/10.1145/3287560.3287574

About the Authors

Zaid Arafat, Assistant Lecturer, Department of Cybersecurity, University of Kerbala (Karbala, 56001 Iraq). E-mail: zaid.q@uokerbala.edu.iq. Scopus Author ID 57963547500, https://orcid.org/0009-0001-0886-5370

Olga V. Yudina, Cand. Sci. (Eng.), Associate Professor, Department of Mathematics and Computer Software, Cherepovets State University (5, Lunacharskogo pr., Cherepovets, 162600 Russia). E-mail: oviudina@chsu.ru. RSCI SPIN-code 7741-5343, https://orcid.org/0009-0005-6367-1076

Zainab A. Abdulazeez, Assistant Lecturer, College of Education for Human Sciences, University of Kerbala (Karbala, 56001 Iraq). E-mail: zainab.abdulhameed@uokerbala.edu.iq. Scopus Author ID 57220186609, https://orcid.org/0009-0004-9801-4888

Об авторах

Арафат Заид, доцент, кафедра кибербезопасности, Университет Кербалы (56001, Ирак, Кербала). E-mail: zaid.q@uokerbala.edu.iq. Scopus Author ID 57963547500, https://orcid.org/0009-0001-0886-5370

Юдина Ольга Вадимовна, к.т.н., доцент, доцент кафедры математического и программного обеспечения ЭВМ, ФГБОУ ВО «Череповецкий государственный университет» (162600, Россия, Череповец, пр-т Луначарского, д. 5). E-mail: oviudina@chsu.ru. SPIN-код РИНЦ 7741-5343, https://orcid.org/0009-0005-6367-1076

Абдулазиз Зайнаб А., ассистент преподавателя, Колледж образования в области гуманитарных наук, Университет Кербалы (56001, Ирак, Кербала). E-mail: zainab.abdulhameed@uokerbala.edu.iq. Scopus Author ID 57220186609, https://orcid.org/0009-0004-9801-4888

The text was submitted by the authors in English
Edited for English language and spelling by Thomas A. Beavitt

Information systems. Computer sciences. Issues of information security
Информационные системы. Информатика. Проблемы информационной безопасности

UDC 004.056.5
https://doi.org/10.32362/2500-316X-2025-13-5-25-40
EDN JKQMQM

RESEARCH ARTICLE

Simulation model of a scalable method for detecting multi-vector attacks taking into account the limitations of computing and information resources of IoT devices

Vyacheslav I. Petrenko, Fariza B. Tebueva, Maxim G. Ogur @, Gennady I. Linets, Valery P. Mochalov

North Caucasus Federal University, Stavropol, 355017 Russia
@ Corresponding author, e-mail: ogur26@gmail.com

• **Submitted:** 14.10.2024 • **Revised:** 13.05.2025 • **Accepted:** 07.08.2025

Abstract

Objectives. The study sets out to develop a scalable method for detecting multi-vector attacks on Internet of Things (IoT) devices. Given the growth of security threats in IoT networks, such a solution must provide high accuracy in detecting attacks with minimal computing costs while taking into account the resource constraints of IoT devices.

Methods. The developed hybrid neural network architecture combines convolutional networks for spatial dependence analysis and long short-term memory networks or gated recurrent units representing types of recurrent neural networks for analyzing time dependencies in network traffic. Model parameters and computational costs are reduced by pruning. A blockchain with a proof of voting[1] consensus mechanism provides secure data management and decentralized verification.

Results. Experiments on the CIC IoT Dataset 2023[2] showed the effectiveness of the model: the accuracy and F1 measure were 99.1%. This confirms the ability to detect known and new attacks in real time with high accuracy and completeness. Processing time is reduced to 12 ms, while memory usage is reduced to 180 MB, which makes the model suitable for devices with limited resources.

Conclusions. The developed model is superior to analogues in terms of accuracy, processing time, and memory usage. Hybrid architecture, pruning, and decentralized verification provide effectiveness against multi-vector IoT threats.

Keywords: multi-vector attacks, Internet of Things, threat detection, neural networks, blockchain, neuronal pruning, cybersecurity, node compromise, consensus, federated learning

[1] Proof of Voting is a consensus algorithm in blockchain networks, in which participants confirm transactions and ensure network security by voting for blocks or transactions.

[2] CIC IoT Dataset 2023. http://cicresearch.ca/IOTDataset/CIC_IOT_Dataset2023/Dataset/. Accessed June 30, 2025.

For citation: Petrenko V.I., Tebueva F.B., Ogur M.G., Linets G.I., Mochalov V.P. Simulation model of a scalable method for detecting multi-vector attacks taking into account the limitations of computing and information resources of IoT devices. *Russian Technological Journal.* 2025;13(5):25–40. https://doi.org/10.32362/2500-316X-2025-13-5-25-40, https://www.elibrary.ru/JKQMQM

Financial disclosure: The authors have no financial or proprietary interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Имитационная модель масштабируемого метода выявления многовекторных атак с учетом ограничений вычислительных и информационных ресурсов IoT-устройств

В.И. Петренко, Ф.Б. Тебуева, М.Г. Огур [@], Г.И. Линец, В.П. Мочалов

Северо-Кавказский федеральный университет, Ставрополь, 355017 Россия
[@] Автор для переписки, e-mail: ogur26@gmail.com

• **Поступила:** 14.10.2024 • **Доработана:** 13.05.2025 • **Принята к опубликованию:** 07.08.2025

Резюме

Цели. Основная цель работы – разработка масштабируемого метода для выявления многовекторных атак на устройства интернета вещей (Internet of Things, IoT). Учитывая рост угроз безопасности в IoT-сетях, решение должно обеспечивать высокую точность обнаружения атак при минимальных вычислительных затратах и с учетом ограничений ресурсов IoT-устройств.

Методы. Для достижения поставленной цели разработана гибридная архитектура нейронных сетей, сочетающая сверточные сети для анализа пространственных зависимостей и сети долгой краткосрочной памяти или Gated Recurrent Units (управляемые рекуррентные блоки) – один из видов рекуррентных нейронных сетей для анализа временных зависимостей в сетевом трафике. Техника обрезки (pruning) сокращает параметры модели и вычислительные затраты. Блокчейн с механизмом консенсуса Proof of Voting[3] обеспечивает безопасное управление данными и децентрализованную верификацию.

Результаты. Эксперименты на датасете CIC IoT Dataset 2023[4] показали эффективность модели: точность и F1-мера составили 99.1%, что подтверждает способность выявлять известные и новые атаки в реальном времени с высокой точностью и полнотой. Время обработки сокращено до 12 мс, использование памяти – до 180 МБ, что делает модель пригодной для устройств с ограниченными ресурсами.

Выводы. Разработанная модель превосходит аналоги по точности, времени обработки и использованию памяти. Гибридная архитектура, обрезка и децентрализованная верификация обеспечивают эффективность против многовекторных угроз IoT. Работа открывает перспективы для исследований в кибербезопасности, предлагая решения для защиты IoT-сетей от сложных атак.

[3] Proof of Voting (алгоритм консенсуса) – это консенсусный алгоритм в блокчейн-сетях, при котором участники подтверждают транзакции и обеспечивают безопасность сети путем голосования за блоки или транзакции. [Proof of Voting is a consensus algorithm in blockchain networks, in which participants confirm transactions and ensure network security by voting for blocks or transactions.]

[4] CIC IoT Dataset 2023. http://cicresearch.ca/IOTDataset/CIC_IOT_Dataset2023/Dataset/. Дата обращения 30.06.2025. / Accessed June 30, 2025.

Ключевые слова: многовекторные атаки, интернет вещей, выявление угроз, нейронные сети, блокчейн, обрезка нейронов, кибербезопасность, компрометация узлов, консенсус, федеративное обучение

Для цитирования: Петренко В.И., Тебуева Ф.Б., Огур М.Г., Линец Г.И., Мочалов В.П. Имитационная модель масштабируемого метода выявления многовекторных атак с учетом ограничений вычислительных и информационных ресурсов IoT-устройств. *Russian Technological Journal*. 2025;13(5):25–40. https://doi.org/10.32362/2500-316X-2025-13-5-25-40, https://www.elibrary.ru/JKQMQM

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

Due to the development of Internet of Things (IoT) technologies networks of IoT devices have become an integral part of modern informational infrastructure. These devices ensure the interaction of numerous systems and platforms in real time to increase the effectiveness, convenience, and flexibility of various sectors: from smart homes and cities to industrial and medical systems. However, IoT devices are often limited in computational and energy resources, making them sensitive for multi-vector cyberattacks. Thus, their wide distribution is associated with an increase in the number of potential threats to information security. Among the most dangerous attacks are DDoS[5], routing attacks, SQL[6] injections, and other forms of multi-vector threats.

Modern methods of detecting attacks, which typically require significant computational resources, can prove inefficient under limited conditions of IoT. This leads to the necessity to develop new approaches that take into account the limitations of computational and informational resources of IoT and simultaneously ensure high security level.

In the present work, we propose a scalable model for detecting multi-vector attacks, which represents a hybrid architecture of Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) or Gated Recurrent Unit (GRM) neural networks (CNN + LSTM/GRU[7]) for analyzing spatiotemporal dependencies of network traffic and decentralized data verification. In order to reduce computational costs, the model incorporates the use of blockchain technologies and neuron pruning. The proposed model, which is oriented towards working in real time given limited resources, is applicable for modern IoT networks. The experimental testing of the efficiency of the developed model using the CIC IoT Dataset 2023 dataset demonstrated its superiority over existing solutions.

[5] Distributed Denial of Service is a form of cyberattack on web systems in order to disable them or make it difficult for ordinary users to access them.

[6] Structured Query Language.

[7] Gated Recurrent Unit.

1. ANALYSIS OF LITERATURE

Sen et al. [1] proposed a mathematical tool for modeling cyberattacks on electric grids involving game theory and the construction of attack graphs. The model is underlaid by the concept of attacker–defender dynamics, where the attacker tries to damage the operation of an electric grid, while the defender tries to prevent the damage using proactive and reactive defense measures. The main advantage of the model is taking into account the attacker–defender dynamics, which makes the model more realistic for use in complex systems. The model uses attack graphs to model multilayer and multistep attacks, taking into account their complexity and diversity. However, the drawbacks of the model include the requirement of initial data on the system and its vulnerabilities, as well as the requirement of the exact evaluation of probabilities of a successful attack and cyber-shutdown cost. This may complicate its practical application for limited data systems.

Lysenko et al. [2] proposed a method for detecting multi-vector cyberattacks on IoT infrastructure by analyzing network traffic and machine learning. Lysenko et al. distinguished four key types of signs helping in accelerated detection of attacks based on data flows, MQTT[8], DNS[9], and HTTP[10]. The method helps increasing the efficiency of detecting attacks by early diagnosis of harmful traffic by analyzing flows and deep analyzing packets for exact detection of multi-vector attacks. This makes this method available for IoT networks with high data volume and complex attack structure. However, the complexity of the method is due to the necessity for an exact determination of a set of signs and their processing, which requires much real-time computational resources in large IoT networks.

Aguru and Erukala [3] proposed a methodology of protecting decentralized IoT networks from multi-vector DDoS attacks using blockchain technologies and deep learning. A Prevent-then-Detect two-stage approach was proposed, where, at the first stage, an intrusion

[8] Message queuing telemetry transport.

[9] Domain name system.

[10] HyperText transfer protocol.

prevention system (IPS) works through a blockchain consortium of validators, while at the second stage, an intrusion detection system (IDS) uses deep learning models for analyzing network traffic and detecting threats. The blockchain ensures security of data transfer between network nodes and controls access to resources of IoT network using intellectual contracts, which are used to determine actions on attack detection and threat prevention. The attack prevention system in a blockchain consortium uses a consensus algorithm to test suspicious traffic. However, the significant computational resources required for the operation of blockchain system and deep neural networks may limit its operability in devices with limited computational possibilities in IoT networks.

Ipole-Adelaiye et al. [4] proposed a method for detecting multi-vector attacks (MVA) using a multilayer perceptron (MLP) approach to analyze network traffic by means of machine learning for detecting various attack vectors. In particular, network data from packet capturing (PCAP) are analyzed to determine anomalous patterns in the behavior of network compounds. The method uses neural networks for data classification and subsequent analysis to increase attack detection accuracy. The MLP, which represents the main component of the proposed system, consists of an input layer, a hidden layer, and an output layer. While the MLP method is suitable for the problems where a high detection accuracy is important, faster models can be used for networks with limited computational resources.

Pakmehr et al. [5] analyze various methods for detecting DDoS attacks on IoT networks with an emphasis on features and challenges that emerge when applying these methods to IoT networks. The authors reviewed several categories of attack detection methods, including signature, anomalous, and hybrid approaches. The mathematical tool includes algorithms based on machine learning, such as Support Vector Machine (SVM), Decision Trees, K-Nearest Neighbors (KNN), and Random Forest. These algorithms are used to classify network traffic and separate anomalous patterns characteristic of DDoS attacks. Although the proposed approaches more efficiently cope with high-volume and various attacks on IoT devices, their efficiency is limited by the complexity of refining the models and the necessity of large computational resources.

Alhakami [6] proposed a mathematical tool for estimating invasion detection methods under Gen V Multi-Vector Attacks. The method is based on a combination of two methods: Fuzzy Analytic Hierarchy Process (Fuzzy AHP) and Technique for Order Preference by Similarity to Ideal Solution (TOPSIS). These methods allow the estimation of various criteria of efficiency of attack detection systems: detection accuracy, adaptivity, scaling, effect on resources, detection time, and automation. A special attention is made to such aspects as adaptation to new threats, the possibility of operating in scaling networks, and the minimization of loading resources at high automation and rapid response.

Saiyed and Al-Anbagi [7] propose an approach for detecting multi-vector DDoS attacks on IoT networks using deep ensemble learning with pruning. Saiyed and Al-Anbagi presented the Deep Ensemble learning with Pruning (DEEPShield) system, which combines the CNN and LSTM networks for analyzing network traffic and detecting both high-volume, and low-volume DDoS attacks. The mathematical tool is based on using an ensemble approach, where CNN extracts spatial signs from network traffic, and LSTM is used to analyze time dependencies. The DEEPShield system demonstrates a high (>90%) accuracy of attack detection and reduces prediction time in comparison with similar models.

Doe et al. [8] describe a hybrid model for analyzing threats and classifying attacks in IoT networks using deep learning and adaptive optimization algorithm Mayfly (LAMOA[11]). The model tends to the detection of routing attacks on IoT networks (sinkhole, wormhole, black hole, and Sybil), which significantly reduce their productivity and security. The model is based on recurrent neural network with long short-term memory for processing time series of network traffic and classification of attacks with the Mayfly adaptive algorithm for optimizing of model hyperparameters. The model, which demonstrates a high ability to exact classification of various types of attacks, is an efficient solution for ensuring security of IoT networks; however, its complexity and computational costs may limit its use in networks with limited sources, thus requiring its further optimization.

Aguru and Erukala [9] propose a lightweight framework for detecting multi-vector DDoS attacks on mobile medicine networks based on IoT using deep learning. Here, the focus on accuracy and efficiency aligns with the purposes of the proposed model and underscores the need for adaptation of detection methods to the specificity of mobile IoT devices, which makes their operation actual for further research in this area.

Petrenko et al. [10] present a method for detecting and counteracting multi-vector threats on decentralized IoT systems that emphasizes the necessity of complex security strategies. The authors underline the importance of integrating various protection methods, including machine learning and blockchain technologies.

The works [11–15] compare approaches to common mitigation of attacks on cloud and fuzzy

[11] Learning-based Adaptive Mayfly Optimization Algorithm.

computations, which may improve the scalability of the model developed in the work, Leng et al. [11] propose methods for improving the scalability and efficiency of protection of IoT networks, which is an important aspect of ensuring security in conditions of growing number of devices and traffic volume. Ali et al. [12] describe a method for protecting decentralized IoT networks from multi-vector DDoS attacks using blockchain technologies and deep learning methods. The proposed two-stage approach combines prevention and detection of attacks, allowing for the efficient control of network threats and an improved level of security. Dalal et al. [13] propose a method for detecting multi-vector attacks based on MLP. Attention is focused on the importance of analysis of network traffic for identifying anomalous patterns, which is a key moment for improving attack detection accuracy. Zahid et al. [14] review various methods for detecting DDoS attacks on IoT networks. The authors discuss signature-, anomalous-, and hybrid approaches in terms of their advantages and disadvantages in the context of IoT. Lungu et al. [15] describe a mathematical tool for estimating invasion detection methods in conditions of fifth-generation multi-vector attacks. By combining decision-making methods, the efficiency of attack detection systems can be estimated, including accuracy and adaptivity.

To construct a simulation model of a scalable method of detection of multi-vector attacks taking into account the limitations of computational and information resources of IoT devices, the following three most suitable models were chosen:

(1) Deep Ensemble Learning with Pruning method [7] using a combination of CNN and LSTM for analyzing network traffic and pruning to reduce computational costs.
(2) Threat Analysis model [8] using LSTM for analyzing routing attacks on IoT networks with adaptive optimization of hyperparameters by the Mayfly algorithm.
(3) Blockchain-based Threat Intelligence Framework method [3] combining blockchain technology with deep learning for protection of IoT networks from multi-vector DDoS attacks.

2. SIMULATION MODEL OF SCALABLE METHOD OF DETECTING MULTI-VECTOR ATTACKS TAKING INTO ACCOUNT LIMITATIONS OF COMPUTATIONAL AND INFORMATIONAL RESOURCES OF IoT DEVICES

The developed model, which should detect multi-vector attacks with high accuracy while minimizing computational costs and taking into account limitations inherent in IoT devices, must be suitable for scaling in large decentralized IoT networks.

The simulation model is built from the following main components:

(1) CNN + LSTM/GRU network traffic analysis module.
(2) Mayfly algorithm for adaptive optimization of hyperparameters.
(3) Proof of Voting (PoV) blockchain-oriented consensus mechanism for decentralized verification.
(4) Neuron pruning for reducing computational costs.

Let us consider these components in more detail.

2.1. CNN + LSTM/GRU network traffic analysis module

The module for network traffic analysis and detection of anomalies in data sequence uses a hybrid architecture of CNN and LSTM (or GRU to reduce computing costs), where:

- The CNN processes spatial signs of network traffic. Input data are represented as a multidimensional tensor, where each element characterizes network packets (e.g., time, size, protocol type). Convolutional layers separate spatial patterns in traffic;
- The LSTM (or GRU) analyzes time dependencies. This helps detect complex multi-vector attacks, which manifests themselves on different time range. Long Short-Term Memory shares information on the previous states of traffic and helps predicting future events, which is important for detecting long-term attacks, such as DDoS.

The network traffic analysis module in the simulation model is based on a hybrid architecture, which combines convolutional neural networks for analysis of spatial dependencies of network traffic and recurrent neuronal networks (LSTM or GRU) for analysis of time dependencies. Such a structure efficiently analyzes multi-vector attacks, which may reveal themselves through complex anomalies in spatial and time dependencies of network traffic.

The convolutional neural network is used to separate spatial signs from network traffic represented by multidimensional data (tensor). The input traffic, including such parameters as time steps, packet size, protocol type, IP addresses, and other metrics, is transformed into a tensor of dimension $\mathbf{X} \in \mathbb{R}^{h \times w \times c}$, where h is the tensor height (number of packets or time steps), w is the tensor width (number of signs or characteristics per packet), and c is the number of channels (e.g., this may be separation by protocols of data types).

The man convolutional equation is written as

$$\mathbf{Y}_{i,j,k} = \sum_{m=1}^{h_k} \sum_{n=1}^{w_k} \mathbf{X}_{i+m,j+n,c} \mathbf{W}_{m,n,k} + b_k, \qquad (1)$$

where $\mathbf{X}_{i,\,j,\,c}$ is the input tensor for position (i, j) in the channel c, $\mathbf{W}_{m,\,n,\,k}$ is the convolutional filter of sizes

$h_k \times w_k$ for the channel k, b_k is the bias for the channel k, and $\mathbf{Y}_{i,j,k}$ is the convolutional result in the channel k.

Convolution (1) is followed by applying an activation function for increasing nonlinearity:

$$\mathbf{Z}_{i,j,k} = \text{ReLU}(\mathbf{Y}_{i,j,k}) = \max(0, \mathbf{Y}_{i,j,k}).$$

Here, ReLU (Rectified Linear Unit) is one of the most popular activation functions, which retains only positive values.

Convolution network separates spatial patterns in network traffic data, such as packet frequency and correlation of various traffic parameters. After spatial signs are isolated using CNN, they are transferred to LSTM to analyze time dependencies. Long Short-Term Memory takes into account the time behavior of traffic and helps isolating multi-vector attacks, which may show through sequential changes in network behavior.

Let us consider the main components of LSTM,

(1) Input gates controls the choice of a new input state to refresh state of memory. The input gates at time t are activated as

$$i_t = \sigma(\mathbf{W}_{\text{in}} \cdot [h_{t-1}, \mathbf{x}_t] + \mathbf{b}_{\text{in}}), \quad (2)$$

where $\mathbf{x}_t$ is the input vector at time t (spatial signs isolated from CNN); h_{t-1} is the hidden state at the previous time step, $\mathbf{W}_{\text{in}}$ is the weight matrix for the input gate, $\mathbf{b}_{\text{in}}$ is the bias vector for the input gate, and σ is a sigmoid normalizing values on the interval [0, 1].

(2) Forgetting gates determine the part of the previous state that should be preserved. The function f_t for activation of forgetting gates can be written as

$$f_t = \sigma(\mathbf{W}_{\text{f}} \cdot [h_{t-1}, \mathbf{x}_t] + \mathbf{b}_{\text{f}}), \quad (3)$$

where $\mathbf{W}_{\text{f}}$ is the weight matrix for the forgetting gate; $\mathbf{b}_{\text{f}}$ is the bias vector for the forgetting gate.

(3) State of memory C_t is refreshed at each time step taking into account new information as:

$$C_t = f_t C_{t-1} + i_t \tanh(\mathbf{W}_{\text{c}} \cdot [h_{t-1}, \mathbf{x}_t] + \mathbf{b}_{\text{c}}), \quad (4)$$

where C_{t-1} is the previous state of memory; C_t is the new state of memory; f_t is the forgetting gates; i_t is the input gates; tanh is the hyperbolic tangent, which is a function of activation used for refreshing state of memory; $\mathbf{W}_{\text{c}}$ is the weight matrix for state of memory; $[h_{t-1}, \mathbf{x}_t]$ is the concatenation of the hidden state at the previous step and the current input vector; and $\mathbf{b}_{\text{c}}$ is the bias vector for refreshing state of memory; the index c indicates the collection of data to memory.

(4) Output gates control the part of state of memory that should be used for refreshing the hidden state. The output gates are activated as follows:

$$o_t = \sigma(\mathbf{W}_{\text{o}} \cdot [h_{t-1}, \mathbf{x}_t] + \mathbf{b}_{\text{o}}), \quad (5)$$

where o_t is the activation function of the output gates, σ is the activation sigmoid, $\mathbf{W}_{\text{o}}$ is the weight matrix for the output gates, and $\mathbf{b}_{\text{o}}$ is the bias vector for the output gates. The new hidden state h_t is calculated as

$$h_t = o_t \tanh(C_t), \quad (6)$$

where h_t is the hidden state at time t, which is used for final classification of network traffic; and C_t is the current state of memory.

Instead of LSTM, GRU can be used, which is modification that uses less computational costs. GRU combines forgotten and input gates into a single refreshing gate, which reduces computational costs and improves model operation under limited resources.

Let us consider the main components of GRU.

(1) Refreshing gates:

$$z_t = \sigma(\mathbf{W}_{\text{z}} \cdot [h_{t-1}, \mathbf{x}_t] + \mathbf{b}_{\text{z}}). \quad (7)$$

Here, z_t is the activation of the refreshing gates, which controls how strongly the current state affects the previous one; the index z shows a refreshing gate (zero).

(2) Removal gates:

$$r_t = \sigma(\mathbf{W}_{\text{r}} \cdot [h_{t-1}, \mathbf{x}_t] + \mathbf{b}_{\text{r}}). \quad (8)$$

Here, r_t is the activation of the removal gates, which controls how strongly the previous state should be forgotten; the index r indicates the reset gate.

(3) Refreshing of the hidden state:

$$h_t = (1 - z_t)\, h_{t-1} + z_t \tanh(\mathbf{W}_{\text{h}}[\mathbf{r}_t h_{t-1}, \mathbf{x}_t] + \mathbf{b}_{\text{h}}), \quad (9)$$

where h_t is the refreshed hidden state at time t, and $\mathbf{W}_{\text{h}}$ is the weight matrix for refreshing the hidden state; the index h displays the gate of hidden state.

Recurrent neural network GRU uses fewer parameters than LSTM, which makes it more suitable for problems requiring smaller computational costs, such as operation under limited resources of IoT devices.

After processing of the CNN and LSTM/GRU data, the model uses a fully connected layer for final classification of traffic. This layer calculates the probabilities that data belongs to one of the classes, e.g., normal traffic or attack:

$$P_{\text{attack}} = \text{Softmax}(\mathbf{W}_{\text{out}} h_{\text{T}} + \mathbf{b}_{\text{out}}), \quad (10)$$

where P_{attack} is the probability that the input traffic is attack, $\mathbf{W}_{\text{out}}$ is the outer layer weights, h_{T} is the hidden state at the last time step, $\mathbf{b}_{\text{out}}$ is the outer layer bias, and Softmax normalizes the outer probability values.

The main variables of the model are the input tensor $\mathbf{X} \in \mathbb{R}^{h \times w \times c}$ of network traffic; the weight matrices $\mathbf{W}$ of layers (CNN, LSTM/GRU); the bias vectors $\mathbf{b}$ of layers (CNN, LSTM/GRU); the input (i_t), forgotten (f_t), and output (o_t) gates to LSTM, respectively; the refreshing (z_t) and removal (r_t) gates to GRU, respectively; the hidden state h_t at time t; the state of memory C_t in LSTM; and the probability P_{attack} that the traffic is attacking.

The network traffic analysis module based on a hybrid architecture of CNN + LSTM (or GRU) combines spatial and time dependencies of network data. Such architecture efficiently detects multi-vector attacks on IoT networks, which is particularly important for systems with limited computational resources.

The model operates as follows.

(1) The input traffic is transformed into a multidimensional tensor, which enters convolutional layers for isolating signs.
(2) The isolated signs enter the LSTM/GRU for analyzing time dependencies.
(3) The model classifies the traffic as normal or attacking.

2.2. Adaptive optimization of hyperparameters using the Mayfly adaptive algorithm

To improve the efficiency of the model and tune it to specific network conditions (e.g., data volume or type of attacks), we use the Mayfly adaptive algorithm [8]. The Mayfly algorithm helps to automatically find the optimal hyperparameters of the model, such as:

- number of layers in CNN and LSTM;
- number of filters and neurons in each layer;
- model learning rate.

The Mayfly adaptive algorithm accelerates model tuning and ensures its optimal productivity without necessity of manual tuning. The algorithm uses evolutionary methods for searching for optimal parameters and is adapted during model learning.

The main steps of the Mayfly adaptive algorithm are the following:

1. Population initialization.
2. Males and females: separation into two groups with different search strategies.
3. Global and local search: search for the best solutions by males and females.
4. Evolution and refreshment of rates and positions.
5. Rendezvous and reproduction.

The main variables and parameters of the algorithm are the following.

N—number of individuals in the population;

x_i^{m}—position of male i in solution space (hyperparameter value);

x_i^{f}—position of female I;

v_i^{m}—male velocity;

v_i^{f}—female velocity;

α, β, γ—male and female motion control coefficients (inertia, acceleration, and interaction, respectively);

g_{best}—global best solution found by all males and females;

p_{best}—personal best solution found by male or female;

λ—attraction coefficient for rendezvous of males and females;

∈—random bias affecting mutation in search.

2.2.1. Population initialization

The Mayfly algorithm begins with random initialization of the initial population of males and females (hyperparameters) in search space. Each male or female is a model hyperparameter vector

$$\mathbf{x}_i = [x_{i1}, x_{i2}, \ldots, x_{id}],$$

where d is the hyperparameter space dimensionality (e.g., number of layers or neurons, training rate, etc.).

The position x_i of each male or female in hyperparameter space is initialized randomly:

$$\mathbf{x}_i^{\mathrm{m}}(0), \mathbf{x}_i^{\mathrm{f}}(0) \sim \mathrm{Uniform}(\mathbf{x}_{\min}, \mathbf{x}_{\max}),$$

where $\mathbf{x}_{\min}$ and $\mathbf{x}_{\max}$ are the hyperparameter space boundaries, and Uniform($\mathbf{x}_{\min}$, $\mathbf{x}_{\max}$) is the function of random sampling from the interval [$\mathbf{x}_{\min}$, $\mathbf{x}_{\max}$].

2.2.2. Refreshment of male velocity and position

Males find solutions in global space, refreshing their position based on personal best solution p_{best} and global best solution g_{best}. The male position refreshment velocity is calculated as:

$$\mathbf{v}_i^{\mathrm{m}}(t+1) = \alpha \mathbf{v}_i^{\mathrm{m}}(t) + \beta_1 r_1 (p_{\mathrm{best},i} - \mathbf{x}_i^{\mathrm{m}}(t)) + \\ + \beta_2 r_2 (g_{\mathrm{best},i} - \mathbf{x}_i^{\mathrm{m}}(t)),$$

where α is the inertia coefficient (which controls how strongly the velocity of the previous step affects the current position); β_1 and β_2 are the acceleration coefficients, which control the effect of the personal and global bets solutions on the velocity refreshment; and r_1 and $r_2 \sim$ Uniform(0, 1) are random values, which ensure random bias in search.

The position of each male is refreshed taking into account its new velocity:

$$\mathbf{x}_i^{\mathrm{m}}(t+1) = \mathbf{x}_i^{\mathrm{m}}(t) + \mathbf{v}_i^{\mathrm{m}}(t+1).$$

2.2.3. Refreshment of female velocity and position

Females perform local search, refreshing their positions based on the distance to males. The female

position refreshment velocity is calculated taking into account the interaction with males:

$$\mathbf{v}_i^{\mathrm{f}}(t+1) = \lambda(x_i^{\mathrm{m}}(t) - x_i^{\mathrm{f}}(t)) + \in,$$

where λ is the attraction coefficient between males and females, and ∈ is a random deviation for ensuring diversity of solutions.

The positions of females is refreshed as follows:

$$\mathbf{x}_i^{\mathrm{f}}(t+1) = \mathbf{x}_i^{\mathrm{f}}(t) + \mathbf{v}_i^{\mathrm{f}}(t+1).$$

2.2.4. Evaluation of solutions

Each male or female is evaluated using a fitness function, which may be related to model, accuracy, learning time, model complexity, and other parameters. The fitness function $F(\mathbf{x}_i)$ for each male or female is calculated as:

$$F(\mathbf{x}_i) = \text{Evaluation Model}(\mathbf{x}_i),$$

where $\mathbf{x}_i$ are the hyperparameters represented by male or female, and Evaluation Model is a function estimating the model productivity at the given hyperparameters.

2.2.5. Rendezvous and reproduction

After the velocities and positions are refreshed, males and females make a rendezvous, which models reproduction in the algorithm. When males and females become close enough, there are crossing over and mutation:

- crossing over transforms part of genetic information (hyperparameters) from males to females:

$$\mathbf{x}_{\text{new}} = \lambda \mathbf{x}_i^{\mathrm{m}} + (1-\lambda)\mathbf{x}_i^{\mathrm{f}},$$

  where $\mathbf{x}_{\text{new}}$ is a new value (position) of male or female as acquired by crossing over, $\mathbf{x}_i^{\mathrm{m}}$ is the current position of male i, $\mathbf{x}_i^{\mathrm{f}}$ is the current position of female i, and λ is a coefficient determining the weight of influence of female in the new state (λ typically ranges from 0 to 1).
- mutation randomly changes some parameters with probability p_{mut}.

2.2.6. Completion criterion

The Mayfly algorithm performs until one of the following conditions is used: the maximum number $T_{\max}$ of iterations has reached, or the fitting function is not improved for several consecutive operations.

2.3. Decentralized verification mechanism based on blockchain technology

The security and reliability of the system under decentralized IoT networks are ensured by PoV blockchain-oriented consensus mechanism [3]. The main functions of this component are the following:

- decentralized verification of attack data, by which several network nodes analyze network traffic and send information on possible attacks to distributed ledger;
- validation of blocks occurs through polling of validating nodes: if more than 50% nodes support attack, the information about it is written in the distributed ledger, and dangerous IP addresses are blocked through action modules.

The decentralized verification module uses blockchain technologies for ensuring data security and preventing attacks in IoT networks. This module operates based on PoV consensus mechanism, which allows network nodes (validators) to vote for data blocks on traffic, attacks, or state of network. Blockchain ensures protection from deception of data, decentralized storage, and automatic performance of such actions as block of harmful IP addresses through action modules.

The main elements and variables are the following.

B—data block containing information on network traffic, detected attacks, or refreshing states of network;

N—number of nodes (validators) in blockchain network;

V_i—voice of validator i for acceptance or rejection of block;

P_{valid}—probability that the block is valid;

T_B—block validation time;

AM—Action Module, which contains data on attacks and blocks of IP addresses.

This subsection presents key steps of processing network traffic and ensuring security in blockchain system. These steps describe how network nodes interact for detecting anomalies, data verification, and automatic blocking of suspicious IP addresses. Every step plays an important rope in creating a reliable and efficient system of protecting from cyberattacks, ensuring data integrity and rapid response to threats.

The steps of processing network traffic and ensuring security in blockchain system are the following:

Step 1. Formation of data block.

Each blockchain network node processes the entering network traffic and, if there is an anomaly or suspicious activity (e.g., multi-vector attack), forms data block B. This block includes the following elements:

$$B = \{\text{Block ID, Data, Previous Hash, Timestamp, Signature}\},$$

where Block ID is the unique block identifier, Data is the information on traffic and possible attacks

(e.g., IP addresses, type of attack, and timestamps), Previous Hash is the hash of the previous block in blockchain for maintaining continuous chain, Timestamp is the block formation time, and Signature is the digital signature of the node that formed the block.

Step 2. PoV consensus mechanism.

After the block is formed, it is transferred to other network nodes for verification using the PoV consensus mechanism.

The block is validated by voting of network nodes in the following order:

- each node analyzes the data block B, tests its integrity and reliability, and then sends its voice V_i (voting may be binary: $V_i = 1$ for acceptance of block and $V_i = 0$ for rejection of block);
- the probability that the block is valid is calculated as

$$P_{\text{valid}} = \frac{\sum_{i=1}^{N} V_i}{N}.$$

If $P_{\text{valid}} \geq 0.5$ (most nodes support the block), then the block is considered valid and is added to the distributed ledger, and if $P_{\text{valid}} < 0.5$, then the block is rejected.

Step 3. Refreshing of distributed ledger

After reaching consensus and supporting block B, it is added to the distributed ledger. Each item in the distributed ledger is related to the previous block through the Previous Hash. Which ensures continuous and invariable data chain. A new block is added to the distributed ledger:

$$B_{\text{new}} = \{\text{Hash}(B_{\text{prev}}), \text{Data}_{\text{new}}, \text{Timestamp}_{\text{new}}, \text{Signature}_{\text{new}}\},$$

where $\text{Hash}(B_{\text{prev}})$ is the hash of the previous block, which guarantees the integrity of the entire chain.

Step 4. Using action modules for automatic blocking of IP addresses.

Blockchain system uses action modules for automatic actions when detecting attack. Action modules Action modules automatically block IP addresses, send notice, and refresh blacklists in network. The structure of action module can have the following form:

$$\text{AM} = \{\text{Source IP (SIP)}, \text{Destination IP (DIP)}, \text{Signature}, \text{Blacklisted IP}, \text{Attack Label}\},$$

where Source IP (SIP) is the IP address, from which traffic enters; Destination IP (DIP) is the IP address of target device; Signature is the digital signature of data for information authentication; Blacklisted IP is the list of IP addresses, which were blocked after detecting attack; and Attack Label is the attack label, which contains type of attack (e.g., DDoS, SQL injection, multi-vector attack).

Step 5. IP address blocking.

Once the data block on attack is confirmed and added to blockchain, action module automatically blocks harmful IP addresses in network. For example, if DDoS traffic is detected, then the IP address of the SIP attacking device is added to the Blacklisted IP blacklist through action module

$$\text{AM(SIP)} = \text{Blacklisted IP}.$$

These data are refreshed at every network node through distributed blockchain structure, which guarantees the consistency of actions of all participants.

Step 6. Block validation time.

For each block B, the block validation time T_B is calculated, which depends on the time t_{vote} of voting of all nodes, the time t_{contract} of performance of all action modules, and the time t_{transmit} of block transmission between nodes:

$$T_B = t_{\text{vote}} + t_{\text{contract}} + t_{\text{transmit}}.$$

The optimization of block validation time is critical for IoT networks with limited resources and high data exchange rate.

2.4. Pruning

Pruning is used to reduce computational costs and optimize model operation at low-power IoT devices. After model learning, minor neurons and their connections are pruned to reduce model volume without significant impairment of its accuracy.

Pruning is performed as follows:

- after learning of neural network, the weights of its connections are analyzed. If the weights are below a given level, then the connections are pruned;
- the model is restarted with a reduced number of neurons and parameters, which reduces its computational complexity and memory requirements.

The main idea is to prune unnecessary or minor neurons or change weights after model learning, insignificantly reducing its productivity.

The main elements and variables are the following:

W—weight matrix of neural network;

b—bias vector of neurons;

$f(\mathbf{W})$—activation function for network weights;

θ—threshold for weight pruning;

M—masking matrix for weight pruning;

n_{total}—total number of parameters (weights) in neural network;

n_{pruned}—number of pruned weights;

p—fraction of pruned weights or neurons.

2.4.1. Determination of significance of weights and neurons

After the neural network is learned, it is necessary to determine, which weights **W** in neural network affect least on output values and can be pruned. This is done by calculating the significance of each weight $W_{i,j}$. As a measure of significance, the metric of absolute value of weight can be used, in which the smaller the weight, the less significant this weight for neuron activation:

$$\text{Significance}(W_{i,j}) = |W_{i,j}|. \quad (11)$$

If the weight is close to zero, then its influence on network output is minimal, and such a weight can be pruned.

2.4.2. Pruning threshold

To determine, which weights should be pruned, threshold θ is introduced. The weights the absolute value of which is smaller than θ are considered minor and removed (equated to zero 0):

$$W_{i,j} = 0 \text{ if } |W_{i,j}| < \theta. \quad (12)$$

The threshold θ is chosen empirically or optimized during experiments. This threshold can be static or dynamic, adapted by analyzing model structure.

2.4.3. Masking matrix

Lest weight pruning affect neurons that have a significant effect on the output values of the model and its productivity, masking matrix **M** is used, denoting which weights should be retained, and which should be zeroed:

$$M_{i,j} = \begin{cases} 1, \text{ if } |W_{i,j}| \geq \theta, \\ 0, \text{ if } |W_{i,j}| < \theta. \end{cases} \quad (13)$$

Pruned masking matrix:

$$\mathbf{W}_{\text{pruned}} = \mathbf{W} \odot \mathbf{M}, \quad (14)$$

where $\odot$ is the elementwise product of the weight matrix **W** and the masking matrix **M**. This guarantees that only significant weights participate in computations, and minor weights are pruned.

2.4.4. Estimation of pruned weight fraction

The fraction of pruned weights or neurons is calculated as

$$p = \frac{n_{\text{pruned}}}{n_{\text{total}}}, \quad (15)$$

where n_{pruned} is the number of pruned weights, i.e., weights for which $|W_{i,j}| < \theta$; and n_{total} is the total number of weights in the model.

2.4.5. Iterative pruning

Simple pruning threshold may be insufficiently efficient for all network layers, especially for deep models with numerous layers. Therefore, iterative pruning may be used: weights are pruned not at once but stepwise, with a gradual increase in threshold θ.

At each iteration, the weights are recalculated using the masking matrix:

$$\mathbf{W}_{\text{new}} = \mathbf{W}_{\text{old}} \odot \mathbf{M}. \quad (16)$$

Then the network is relearned on new data to restore its accuracy after pruning. This process is repeated several times until the pruned weight fraction reaches the desired level p.

2.4.6. The main steps of the developed simulation model and metric for estimating the model quality after weight pruning

The main steps of the model operation are the following:

(1) traffic analysis using CNN + LSTM for detecting spatiotemporal signs and identifying anomalies;
(2) model optimization: the Mayfly algorithm automatically tunes model hyperparameters depending on the conditions of network and data, which ensures its adaptability;
(3) suspicious IP addresses and traffic are verified and blocked through blockchain consortium. If the attack on IP addresses is confirmed, they are blocked through action modules;
(4) neuron pruning: after the initial learning and verification of the model, neurons are pruned to reduce computational complexity and model adaptation to resources of IoT devices.

After pruning, it is important to estimate the changes in model productivity and resource intensity. The main metrics of model quality after weight pruning are the following:

(1) the fraction of accurate predictions among the total amount of predictions (accuracy of correct predictions):

$$E_1 = \text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \quad (17)$$

where TP is the number of true positive predictions (correct predictions of attacks), TN is the number of true negative predictions (correct predictions of normal traffic), FP is the number of false positive predictions (false alerts), and FN is the number of false negative predictions (lost attacks);

(2) recall:

$$E_2 = \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (18)$$

which determines the model ability to detect all attacks in a sample;

(3) precision (fraction of true positive predictions among all positive predictions (accuracy of true positive predictions)):

$$E_3 = \text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (19)$$

which determined how much model predictions of positive classes are correct;

(4) F1 metric:

$$E_4 = 2\frac{E_2 E_3}{E_2 + E_3}, \quad (20)$$

which is a harmonic mean between recall and precision;

(5) time of data processing and performing model predictions:

$$E_5 = T_{\text{pruned}} = T_{\text{original}}(1 - p), \quad (21)$$

where T_{pruned} is the computational time after pruning, T_{original} is the computational time by the initial model, and p is the fraction of pruned neurons;

(6) memory usage–(reduction of memory usage after weight pruning):

$$E_6 = M_{\text{pruned}} = M_{\text{original}}(1 - p), \quad (22)$$

where M_{pruned} is the memory required for storing pruned model, and M_{original} is the memory necessary for storing the initial model.

3. IMPLEMENTATION AND EXPERIMENT

3.1. Design of experiment

In this section, we experimentally tested the developed simulation model of detecting multi-vector attacks taking into account the limitations of computational and informational resources of IoT devices.

The experiment uses CIC IoT Dataset 2023, which contains data of network traffics, both normal, and attacking, with various characteristics and signs, such as:

- IP addresses (Source/Destination);
- ports (Source/Destination);
- connection time;
- packet size;
- protocols (TCP[12], UDP[13], HTTP, DNS);
- attack mark (e.g., DDoS, SQL injection, Brute Force).

[12] Transmission Control Protocol.
[13] User Datagram Protocol.

The dataset is divided into several classes:

- normal traffic;
- attacking traffic (various types of attacks).

The experiment compares the results of the developed simulation model with several similar methods used to detect attacks on IoT networks.

The CIC IoT Dataset 2023 data were normalized using Min-Max Scaling to reduce all signs to the range [0, 1]. Outliers were removed by Interquartile Range (IQR) method, and new signs were generated by aggregation of time characteristics of traffic (e.g., mean number of packets in 10 s).

The testing was performed on a workstation based on Macbook Pro notebook (Apple Inc., USA) with M2 Pro processor (includes 12 processor cores (8 performance cores and 4 efficiency cores), 19 graphics engines, and 16-core neural coprocessor) and16-GB random access memory with a throughput efficiency of about 200 GB/s.

3.2. Structure of experiment

The experiment was performed in several steps:

(1) Data preparation, at which the CIC IoT Dataset 2023 data are divided into learning and testing samplings in the ratio 70 : 30. Then, the data are preliminarily processed: signs are normalized, outliers are removed, and new signs are generated (if necessary).

(2) Model learning, at which the proposed model is learned based on the hydride architecture CNN + LSTM/GRU using neuron pruning to reduce computational costs.

The model hyperparameters are optimized using the Mayfly algorithm.

For comparison, other models are also learned, such as Random Forest, SVM, Deep Learning (MLP).

The processing time of a single data packet is found as the sum of convolution time (1), LSTM/GRU processing time (2)–(9), and classification time (10). Owing to neuron pruning (11)–(16), the model significantly reduces the number of parameters, which decreases the computational costs and improves productivity on devices with limited resources. This makes it possible to efficiently use the model under real conditions, such as real-time monitoring systems, where fast response and minimum memory consumption are important. The quality of the simulation model is estimated using metrics (17)–(22).

3.3. Analysis of the obtained results

The efficiency of the proposed model was estimated in experimental studies. Table estimates the efficiency using evaluation metrics (17)–(22) of various models, including the proposed model based on the CNN + LSTM/GRU hybrid architecture using neuron pruning.

Table. Experimental results

Model	Evaluation metrics					
	E_1, %	E_2, %	E_3, %	E_4, %	E_5, ms	E_6, MB
Random Forest	96.5	95.7	97.1	96.4	35	220
SVM	94.3	92.6	94.5	93.5	50	250
Deep Learning (MLP)	97.8	97.2	98.0	97.6	20	210
Simulation model CNN + LSTM/GRU	99.1	99.3	98.9	99.1	12	180

Figure 1 compares the results of operation of the proposed simulation model and its analogs in accordance with evaluation metrics (17)–(20). Figure 1 shows that the CNN + LSTM/GRU model significantly exceeds the other models in all presented metrics, which confirms its high efficiency in detecting multi-vector attacks.

Figure 2 compares the results of operation of the proposed simulation model and its analogs in accordance with evaluation metrics E_5 (21). Figure 2 shows that the model CNN + LSTM/GRU has the minimum data processing time and the minimum prediction performance time (12 ms), which makes it particularly suitable for real-time use, whereas the other models require much more time.

Figure 3 compares the results of operation of the proposed simulation model and its analogs in accordance with quality metric E_6 (22). Figure 5 shows that the model CNN + LSTM/GRU requires the minimum memory volume (180 MB) for processing the input data of 1 million examples with 10 signs, which makes it more efficient for using on devices with limited computational resources in comparison with the other models, such as Random Forest and SVM.

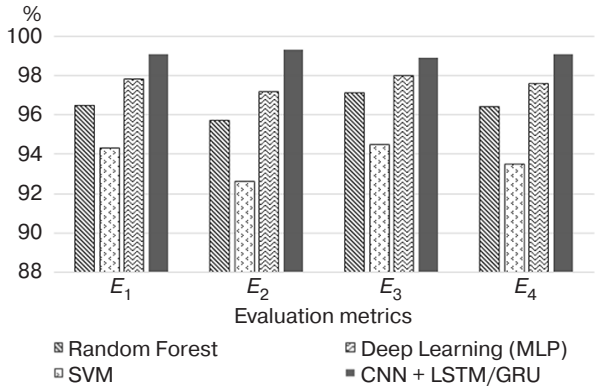


Fig. 1. Comparison of the results of operation of the developed simulation model CNN + LSTM/GRU and its analogues in accordance with metrics E_1 (17) – E_4 (20)

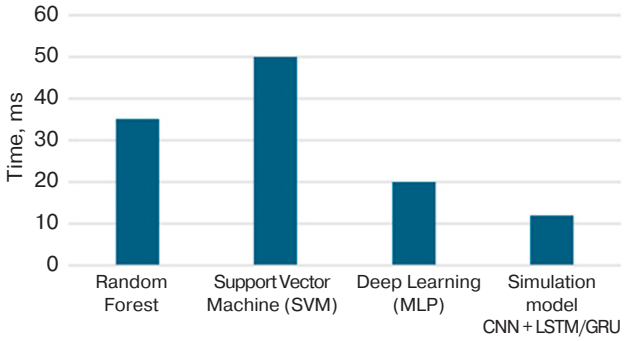


Fig. 2. Comparison of the results of operation of the developed simulation model CNN + LSTM/GRU and its analogues in accordance with metric E_5 (21)

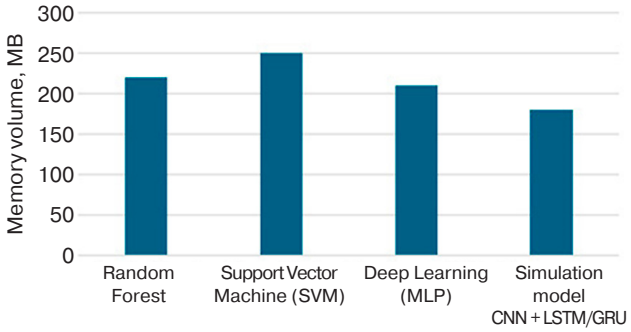


Fig. 3. Comparison of the results of operation of the developed simulation model CNN + LSTM/GRU and its analogues in accordance with metric E_6 (22)

The performed experiment confirms the following:

(1) The developed simulation model attained high accuracy of detecting attacks at a level of 99.1%, which suggests its ability to efficiently identify both known and new types of attacks in real time. Thus, the proposed architecture based on the CNN + LSTM/GRU hybrid model can be successfully applied in the context of analyzing network traffic.

(2) The F1 metric of the developed simulation model is 99.1%, which indicates high balance between precision (19) and recall (18). This means that the model not only accurately identifies attacks, but also minimizes the number of false alarms and miss.

(3) The request processing time in the proposed simulation model at hardware resources stated in 3.1 of this work is reduced to 12 ms, making this model particularly useful for systems requiring fast response, such as real-time monitoring systems. This represents a significant advantage over the other models, which require more processing time.

(4) The developed simulation model uses only 180 MB memory, making it suitable for implementing on devices with limited computational resources. This is particularly important for IoT devices, which are often limited in memory and computational power.

(5) Neuron pruning significantly reduced the number of model parameters from 1.5 million to 300 thousand, which, in turn, decreased computational costs by 80% and improved productivity. This confirms that model optimization approaches contribute significantly to its successful use under limited resources.

CONCLUSIONS

In this work, we have proposed a simulation model of a scalable method for detecting multi-vector attacks on IoT devices that takes into account the limitations of computational and informational resources. The creation of an efficient solution capable of detecting attacks with high accuracy is a key objective given the growing security threats in IoT.

The proposed model is based on a hybrid architecture of neural networks that combines convolutional neural networks CNN for analyzing spatial dependencies and long short-term memory networks LSTM for analyzing time dependencies of network traffic. An important aspect is pruning, which significantly reduces the number of model parameters to decrease computational costs. The use of blockchain technologies with a PoV consensus mechanism ensures data security and decentralized verification, which is critically important for protecting IoT networks from multi-vector attacks.

The experimental testing using the CIC IoT Dataset 2023 dataset demonstrated the high efficiency of the proposed model. The achieved attack detection accuracy 99.1% confirms its ability to exactly identify both known and new types of attacks in real time. The F1 metric of 99.1% indicates a balance between precision and recall, which is critically important for cybersecurity systems in which both false alarms and unidentified attacks should be minimized. In addition to high accuracy, the request processing time was reduced to 12 ms. This allows the model to function efficiently under rapid response conditions such as real-time monitoring systems. Memory use was also optimized to only 180 MB, which makes it suitable for devices with limited computational resources.

Thus, the developed simulation model exceeds the existing solutions in key metrics, such as precision, processing time, and memory use. The high efficiency of the model during multi-vector threats to IoT is ensured by its hybrid architecture, neuron pruning, and decentralized verification.

This work opens new horizons for further research in cybersecurity that will lead to efficient solutions for protecting IoT networks from complex cyberthreats. Future studies should aim to integrate additional machine learning and deep learning methods for increasing the accuracy and stability of the model to new types of attacks. It is also worth considering the possibility of optimizing algorithms to reduce computational costs and increase data processing rates. Under current conditions of increased device numbers

and traffic volumes, it becomes crucial to continue to improve the scalability and stability of blockchain-oriented solutions.

ACKNOWLEDGMENTS

This research was supported by the Information Security grant from the Moscow Technical University of Communications and Informatics (MTUSI) (agreement No. 40469/17-23-K, "Development of an intrusion detection method using multi-vector attack scenarios in a decentralized IoT environment."

Authors' contributions

V.I. Petrenko—research idea, planning the study, and scientific editing the article.

F.B. Tebueva—research idea, planning the study, and scientific editing the article.

M.G. Ogur—conducting the research, performing the experimental part of the work, analysis of the obtained data, formulating the results, and writing the text of the article.

G.I. Linets—consultations on conducting the research and scientific editing the article.

V.P. Mochalov—consultations on conducting the research and scientific editing the article.

REFERENCES

1. Sen Ö., Ivanov B., Henze M., Ulbig A. Investigation of Multi-stage Attacks and Defense Modeling for Data Synthesis. In: *Proceedings of the International Conference on Smart Energy Systems and Technologies* (*SEST*). IEEE; 2023. P. 1–12. https://doi.org/10.1109/SEST57387.2023.10257329
2. Lysenko S., Bobrovnikova K., Kharchenko V., Savenko O. IoT Multi-Vector Cyberattack Detection Based on Machine Learning Algorithms: Traffic Features Analysis, Experiments, and Efficiency. *Algorithms*. 2022;15(7):239. https://doi.org/10.3390/a15070239
3. Aguru A., Erukala S. OTI-IoT: A Blockchain-based Operational Threat Intelligence Framework for Multi-vector DDoS Attacks. *ACM Trans. Internet Technol*. 2024;24(3):15.1–15.31. https://doi.org/10.1145/3664287
4. Ipole-Adelaiye N., Tatama F.B., Egena O., Jenom M., Ibrahim L. Detecting Multi-Vector Attack Threats Using Multilayer Perceptron Network. *IRE Journals*. 2024;8(1):119–123.
5. Pakmehr A., Aßmuth A., Taheri N., Ghaffari A. DDoS attack detection techniques in IoT networks: a survey. *Cluster Comput*. 2024;27(4):14637–14668. https://doi.org/10.1007/s10586-024-04662-6
6. Alhakami W. Evaluating modern intrusion detection methods in the face of Gen V multi-vector attacks with fuzzy AHP-TOPSIS. *PLoS One*. 2024;19(5):e0302559. https://doi.org/10.1371/journal.pone.0302559
7. Saiyed M.F., Al-Anbagi I. Deep Ensemble Learning With Pruning for DDoS Attack Detection in IoT Networks. *IEEE Trans. Machine Learning Commun. Networks.* 2024;2:596–616. https://doi.org/10.1109/TMLCN.2024.3395419
8. Liebl S. *Threat Modelling for Internet of Things Devices.* Research Report 2023 of the Technical University OTH Amberg-Weiden. 2023. Available from URL: https://www.researchgate.net/publication/369488078. Accessed February 25, 2025.
9. Aguru A.D., Erukala S.B. A lightweight multi-vector DDoS detection framework for IoT-enabled mobile health informatics systems using deep learning. *Inf. Sci*. 2024;662:120209. https://doi.org/10.1016/j.ins.2024.120209
10. Petrenko V.I., Tebueva F.B., Ogur M.G., Linets G.I., Mochalov V.P. Methodology for detecting and countering multi-vector threats to information security of a decentralized IoT system. *Int. J. Open Inf. Technol*. 2025;13(1):13–24 (in Russ.).
11. Leng S., Guo Y., Zhang L., Hao F., Cao X., Li F., Kou W. Online and Collaboratively Mitigating Multi-Vector DDoS Attacks for Cloud-Edge Computing. In: *ICC 2024 – International Conference on Communications*. 2024. P. 1394–1399. https://doi.org/10.1109/ICC51166.2024.10623052
12. Ali M., Saleem Y., Hina S., Shah G.A. DDoSViT: IoT DDoS attack detection for fortifying firmware Over-The-Air (OTA) updates using vision transformer. *Internet of Things*. 2025;30:101527. https://doi.org/10.1016/j.iot.2025.101527
13. Dalal S., Lilhore U.K., Faujdar N., Simaiya S., et al. Next-generation cyberattack prediction for IoT systems: leveraging multi-class SVM and optimized CHAID decision tree. *J. Cloud Comput*. 2023;12:137. https://doi.org/10.1186/s13677-023-00517-4

14. Zahid F., Funchal G., Melo V., Kuo M.M.Y., et al. DDoS attacks on smart manufacturing systems: A cross-domain taxonomy and attack vectors. In: *2022 20th IEEE International Conference on Industrial Informatics* (*INDIN*). 2022. P. 214–219. https://doi.org/10.1109/INDIN51773.2022.9976172
15. Lungu N., Dash B.B., De U.C., Dash B.B., et al. Multi-vector Monitoring, Detecting and Classifying GPU Side-Channel Attack Vectors on a Secure GPU Execution Framework. In: *2024 8th International Conference on I-SMAC* (*IoT in Social, Mobile, Analytics and Cloud*). 2024. P. 500–505. https://doi.org/10.1109/I-SMAC61858.2024.10714895

About the Authors

Vyacheslav I. Petrenko, Cand. Sci. (Eng.), Associate Professor, Head of the Department of Organization and Technology of Information Security, Prof. Nikolay Chervyakov Faculty of Mathematics and Computer Sciences, North-Caucasus Federal University (1, Pushkina ul., Stavropol, 355017 Russia). E-mail: vipetrenko@ncfu.ru. Scopus Author ID 57189512011, ResearcherID A-3196-2017, RSCI SPIN-code 3923-4295, https://orcid.org/0000-0003-4293-7013

Fariza B. Tebueva, Dr. Sci. (Phys.-Math.), Associate Professor, Professor, Department of Computational Mathematics and Cybernetics, Prof. Nikolay Chervyakov Faculty of Mathematics and Computer Sciences, North-Caucasus Federal University (1, Pushkina ul., Stavropol, 355017 Russia). E-mail: ftebueva@ncfu.ru. Scopus Author ID 57189512319, ResearcherID H-4548-2017, RSCI SPIN-code 9343-7504, https://orcid.org/0000-0002-7373-4692

Maxim G. Ogur, Senior Lecturer, Department of Computational Mathematics and Cybernetics, Prof. Nikolay Chervyakov Faculty of Mathematics and Computer Sciences, North-Caucasus Federal University (1, Pushkina ul., Stavropol, 355017 Russia). E-mail: ogur26@gmail.com. ResearcherID B-1332-2017, RSCI SPIN-code 7180-6971, https://orcid.org/0000-0002-2387-0901

Gennady I. Linets, Dr. Sci. (Eng.), Professor, Department of Digital, Robotic Systems and Electronics, Institute of Advanced Engineering, North-Caucasus Federal University (1, Pushkina ul., Stavropol, 355017 Russia). E-mail: kbytw@mail.ru. Scopus Author ID 6506372022, RSCI SPIN-code 1452-6823, https://orcid.org/0000-0002-2279-3887

Valery P. Mochalov, Dr. Sci. (Eng.), Professor, Department of Digital, Robotic Systems and Electronics, Institute of Advanced Engineering, North-Caucasus Federal University (1, Pushkina ul., Stavropol, 355017 Russia). E-mail: mochalov.valery2015@yandex.ru. Scopus Author ID 57202300745, RSCI SPIN-code 8695-1648, https://orcid.org/0000-0002-5131-5649

Об авторах

Петренко Вячеслав Иванович, к.т.н., доцент, заведующий кафедрой организации и технологии защиты информации, факультет математики и компьютерных наук имени профессора Н.И. Червякова, ФГАОУ ВО «Северо-Кавказский федеральный университет» (355017, Россия, Ставрополь, ул. Пушкина, д. 1). E-mail: vipetrenko@ncfu.ru. Scopus Author ID 57189512011, ResearcherID A-3196-2017, SPIN-код РИНЦ 3923-4295, https://orcid.org/0000-0003-4293-7013

Тебуева Фариза Биляловна, д.ф.-м.н., доцент, профессор кафедры вычислительной математики и кибернетики, факультет математики и компьютерных наук имени профессора Н.И. Червякова, ФГАОУ ВО «Северо-Кавказский федеральный университет» (355017, Россия, Ставрополь, ул. Пушкина, д. 1). E-mail: ftebueva@ncfu.ru. Scopus Author ID 57189512319, ResearcherID H-4548-2017, SPIN-код РИНЦ 9343-7504, https://orcid.org/0000-0002-7373-4692

Огур Максим Геннадьевич, старший преподаватель, кафедра вычислительной математики и кибернетики, факультет математики и компьютерных наук имени профессора Н.И. Червякова, ФГАОУ ВО «Северо-Кавказский федеральный университет» (355017, Россия, Ставрополь, ул. Пушкина, д. 1). E-mail: ogur26@gmail.com. ResearcherID B-1332-2017, SPIN-код РИНЦ 7180-6971, https://orcid.org/0000-0002-2387-0901

Линец Геннадий Иванович, д.т.н., профессор, профессор департамента цифровых, робототехнических систем и электроники, институт перспективной инженерии, ФГАОУ ВО «Северо-Кавказский федеральный университет» (355017, Россия, Ставрополь, ул. Пушкина, д. 1). E-mail: kbytw@mail.ru. Scopus Author ID 6506372022, SPIN-код РИНЦ 1452-6823, https://orcid.org/0000-0002-2279-3887

Мочалов Валерий Петрович, д.т.н., профессор, профессор департамента цифровых, робототехнических систем и электроники, институт перспективной инженерии, ФГАОУ ВО «Северо-Кавказский федеральный университет» (355017, Россия, Ставрополь, ул. Пушкина, д. 1). E-mail: mochalov.valery2015@yandex.ru. Scopus Author ID 57202300745, SPIN-код РИНЦ 8695-1648, https://orcid.org/0000-0002-5131-5649

Translated from Russian into English by Vladislav Glyanchenko
Edited for English language and spelling by Thomas A. Beavitt

Information systems. Computer sciences. Issues of information security
Информационные системы. Информатика. Проблемы информационной безопасности

UDC 519.95:621.3
https://doi.org/10.32362/2500-316X-2025-13-5-41-50
EDN KVUAJF

RESEARCH ARTICLE

Analysis of information transmission processes in multimode fiber-optic networks with a token-based access method

Dmitry V. Zhmatov [@], Alexander S. Leontyev

MIREA – Russian Technological University, Moscow, 119454 Russia
[@] Corresponding author, e-mail: zhmatov@mirea.ru

• **Submitted:** 10.02.2025 • **Revised:** 17.03.2025 • **Accepted:** 22.07.2025

Abstract

Objectives. The study sets out to develop and analyze a mathematical model for information transmission in multimode fiber-optic ring networks using a token-based access method to ensure efficient interaction between Internet of Things (IoT) devices. The work aims to evaluate the probabilistic and time-related characteristics, as well as the reliability and performance of the network infrastructure to optimize data transmission parameters, taking into account the specifics of IoT and the peculiarities of the fiber-optic medium.

Methods. Reliability theory methods are used to assess the network's resilience to failures and increase its operational efficiency, along with techniques from the theory of stochastic processes to model the dynamics of data transmission under varying loads and approaches from queueing theory to analyze traffic distribution and packet queue management. The Laplace–Stieltjes transform is applied to derive functional equations that describe the probabilistic and time-related data transmission characteristics, enabling precise mathematical modeling of network processes.

Results. The information transmission processes occurring in multimode fiber-optic networks with token access in the context of IoT systems were studied. The temporal characteristics of packet transmission for different classes, including critical IoT device data, were analyzed.

Conclusions. The results confirm that multimode fiber-optic media provide an efficient foundation for IoT infrastructure that offers both high throughput and fault tolerance. By incorporating reliability characteristics into the model, it was possible to account for the impact of fiber-optic medium and network node failures on performance. Optimizing the parameters of the token-based access method, including time intervals and token transmission policies, significantly improves overall network performance by reducing collision probability and increasing throughput. The developed mathematical model provides an effective tool for analyzing and designing local networks based on multimode fiber-optic technologies. This fact is especially important for networks serving critical infrastructure.

Keywords: FDDI networks, token-based access method, models, temporal characteristics, failures, performance

For citation: Zhmatov D.V., Leontyev A.S. Analysis of information transmission processes in multimode fiber-optic networks with a token-based access method. *Russian Technological Journal.* 2025;13(5):41–50. https://doi.org/10.32362/2500-316X-2025-13-5-41-50, https://www.elibrary.ru/KVUAJF

Financial disclosure: The authors have no financial or proprietary interest in any material or method mentioned.
The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Анализ процессов передачи информации в многомодовых оптоволоконных сетях с маркерным методом доступа

Д.В. Жматов [@], А.С. Леонтьев

МИРЭА – Российский технологический университет, Москва, 119454 Россия
[@] Автор для переписки, e-mail: zhmatov@mirea.ru

• **Поступила:** 10.02.2025 • **Доработана:** 17.03.2025 • **Принята к опубликованию:** 22.07.2025

Резюме

Цели. Целью работы являются разработка и анализ математической модели передачи информации в многомодовых оптоволоконных кольцевых сетях с маркерным методом доступа для обеспечения эффективного взаимодействия устройств интернета вещей (Internet of Things, IoT). Работа направлена на оценку вероятностно-временных характеристик, надежности и производительности сетевой инфраструктуры, а также оптимизацию параметров передачи данных с учетом специфики IoT и особенностей оптоволоконной среды.

Методы. В ходе исследования применены методы теории надежности для оценки устойчивости сети к отказам и увеличения ее эксплуатационной эффективности, методы теории случайных процессов для моделирования динамики передачи данных в условиях изменяющейся нагрузки, а также методы теории массового обслуживания для анализа распределения трафика и управления очередями пакетов. Дополнительно использовано преобразование Лапласа – Стилтьеса, позволяющее вывести функциональные уравнения, описывающие вероятностно-временные характеристики передачи данных и обеспечивающие точное математическое моделирование сетевых процессов.

Результаты. Исследованы процессы передачи информации в многомодовых оптоволоконных сетях с маркерным доступом в контексте IoT-систем. Проведен анализ временных характеристик передачи пакетов различных классов, включая критически важные данные IoT-устройств.

Выводы. Результаты исследования подтверждают, что многомодовая оптоволоконная среда является эффективной основой для IoT-инфраструктуры, обеспечивая высокую пропускную способность и устойчивость к отказам. Включение характеристик надежности в модель позволило учесть влияние отказов оптоволоконной среды и узлов сети на производительность. Оптимизация параметров маркерного метода доступа, включая временные интервалы и политику передачи маркеров, существенно повышает общую производительность сети, снижая вероятность коллизий, и увеличивает пропускную способность. Разработанная математическая модель предоставляет эффективный инструмент для анализа и проектирования локальных сетей на основе многомодовых оптоволоконных технологий. Это особенно важно для сетей, обслуживающих критически важные инфраструктуры.

Ключевые слова: FDDI-сети, маркерный метод доступа, модели, временные характеристики, отказы, производительность

Для цитирования: Жматов Д.В., Леонтьев А.С. Анализ процессов передачи информации в многомодовых оптоволоконных сетях с маркерным методом доступа. *Russian Technological Journal*. 2025;13(5):41–50. https://doi.org/10.32362/2500-316X-2025-13-5-41-50, https://www.elibrary.ru/KVUAJF

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

Multimode access methods are a type of data transmission technique in fiber optic systems that are widely used to increase the capacity and efficiency of information transmission. The basic principle consists in the simultaneous transmission of light signals through several independent modes representing different trajectories of light propagation inside the fiber core. Since each mode acts as an independent communication channel, this enables parallel data transmission.

Especially in high-bandwidth systems, multimode access methods enable efficient scaling of transmitted data volumes. With an increase in the number of modes or the bandwidth of each mode, such systems find application in high-speed optical networks, including backbone communication channels and local networks.

The characteristics of data processing and transmission systems in network structures are largely determined by the reliability of the transmission medium, taking into account possible failures and malfunctions [1–3]. However, in many studies, the temporal parameters of transmission are considered in isolation, without taking into account reliability factors, while fault tolerance analysis often does not include the influence of information processing and transmission technologies. Therefore, when studying multimode fiber optic transmission systems, mathematical methods that consider possible failures of the transmission medium are used to evaluate the temporal characteristics. In [1, 2], the authors developed mathematical methods for analyzing the probabilistic-temporal characteristics of data transmission in networks with a token access method, taking into account the impact of failures, as well as methods for assessing the resource loading and performance of such network structures. Due to their versatility and potential, multimode access methods are becoming an important tool for building high-speed optical networks, including backbone communication channels, data centers and local area networks [4–6]. These technologies provide flexibility of scaling by increasing the number of modes, improving transmission quality, and increasing the throughput of each modulation trajectory.

1. MULTIMODE FIBER STRUCTURE

In a multimode fiber optic cable, the core diameter is larger compared to single-mode fibers. This fact enables light to propagate through different modes. Each mode represents a specific path of light propagation within the fiber, determined by physical laws (light entry angle, core and cladding properties).

In multimode method, data is distributed among the available modes. Each mode acts as a separate communication channel, providing independent data transmission [7, 8]. Such transmission methods utilize channel sharing features including:

- wavelength-division multiplexing (WDM), where each mode is allocated a specific wavelength;
- time division multiplexing (TDM), where each mode transmits data in the allotted time slot;
- increase in throughput capacity without significant increase in equipment cost;
- scalability of the system by increasing the number of modes;
- parallel data processing, which is especially important for highly loaded networks [9–11].

Since different modes pass through the fiber at different speeds, temporal blurring of the signal can occur. Multimode fibers are typically used over short distances (e.g., in local area networks) due to high signal loss and dispersion.

Increasing the complexity of signal processing implies the need for specialized equipment for separating modes, eliminating mutual interference between them, and ensuring the stability of data transmission, which requires additional tuning and calibration of the system to achieve optimal signal quality.

In the context of the industrial internet, multimode fiber optic systems are used to provide high-speed data exchange in the presence of a large number of connected devices[1] [12–17].

Multimode systems can be easily integrated with a variety of transmission protocols including Ethernet and Fiber Channel[2]. Increasing the number of modes or upgrading the system can increase network performance at no significant cost.

2. LIMITATIONS OF THE MULTIMODE METHOD

Data transmission in the network is carried out considering packet length requirements, data flow intensity from each node, and time constraints specified by the customer.

To investigate the process of data transmission in the network, the token access method was used. Modeled network parameters were: number of nodes N_{nodes}; intensity of packets arrival in the network nodes for each mode λ_k, $k = \overline{1, N}$; bandwidth of the transmitting medium C; failure rate of the transmitting medium λ_{fail}; length of the transmitted packet L_{pack}; limitation on the

[1] Kleinrock L. To Mario Gerla, the Maestro of Networks. *Ad Hoc Networks*. 2019;88:178–179. https://www.lk.cs.ucla.edu/data/files/Mario%20Gerla%20tribute%20by%20Len.pdf. Accessed June 24, 2025.

[2] Fiber Channel is a family of protocols for high-speed data transmission. The protocols are standardized by the T11 Technical Committee, part of the International Committee for Information Technology Standards.

time of packet transmission of each mode in the network $T_{\text{dir}}^{(1)}$ (directive time). The following parameters were determined for each mode: node polling cycle $Z^{(1)}$; node utilization ρ_k, $k=\overline{1,N}$ and transmission medium R_{medium}; packet delivery time $T_k^{(1)}$; probability of timely packet delivery taking into account the reliability of the transmission medium Q_k, $k=\overline{1,N}$; network performance λ_{tot}.

The total intensity of timely served flow for each mode is chosen as an estimate of network performance:

$$\lambda_{\text{tot}}=\sum_{k=1}^{N}\lambda_k Q_k.$$

Let the bandwidth C, the network architecture with N_{nodes} nodes, N packet flows for each mode with intensities λ_k, $k=\overline{1,N}$, and data transmission time constraints be given.

The packet flows λ_k for each mode and the transmission medium failures λ_{fail} are Poissonian in nature; the buffer devices of the nodes have unlimited capacity. The number of nodes in the network and the number of packet flows are the same for each mode.

3. THEOREM ON THE EVALUATION OF THE POLLING CYCLE IN TOKEN NETWORKS

The polling cycle of network nodes for each mode is directly proportional to the number of nodes, length of a marker, and inversely proportional to the throughput capacity of the transmitting medium and the probability that the medium is free from packet transmission of this mode.

If these conditions of the theorem are fulfilled, the cycle of node polling in the reliable operation mode is described by the following relation:

$$Z^{(1)}=\frac{N_{\text{node}}\dfrac{L_{\text{m}}}{C}}{1-\sum\limits_{k=1}^{N}\lambda_k\dfrac{L_{\text{pack}}}{C}},\qquad(1)$$

where L_{m} is the length of the marker, and taking into account failures of the transmitting medium—by the formula:

$$Z_{\text{fail}}^{(1)}=\frac{Z^{(1)}}{1-\lambda_{\text{fail}}F_{\text{fail}}^{(1)}},\qquad(2)$$

where $F_{\text{fail}}^{(1)}$ is the average time of medium performance recovery after failure.

Proof. Let $T_{\text{pack}}=\dfrac{L_{\text{pack}}}{C}$ is the time of the packet transmission from one node to another. Obviously, the polling cycle of the network nodes (the average interval between two consecutive polls of a node) is equal to:

$$Z^{(1)}=\sum_{k=1}^{N}\left\{\rho_k\left(\frac{L_{\text{pack}}}{C}+\frac{L_{\text{m}}}{C}\right)+(1-\rho_k)\frac{L_{\text{m}}}{C}\right\}=\sum_{k=1}^{N}\left\{\rho_k\frac{L_{\text{pack}}}{C}+\frac{L_{\text{m}}}{C}\right\}.\qquad(3)$$

In the stationary mode, $\lambda_k Z^{(1)}=\rho_k$. Multiplying Eq. (3) by λ_n, $n=\overline{1,N}$, we obtain a system of N linear inhomogeneous equations relative to ρ_n:

$$\lambda_n\sum_{k=1}^{N}\left\{\rho_k\left(\frac{L_{\text{pack}}}{C}\right)+\frac{L_{\text{m}}}{C}\right\}=\rho_n.\qquad(4)$$

Let us choose any two equations for ρ_i and ρ_j:

$$\lambda_i\sum_{k=1}^{N}\left\{\rho_k\frac{L_{\text{pack}}}{C}+\frac{L_{\text{m}}}{C}\right\}=\rho_i,\qquad(5)$$

$$\lambda_j\sum_{k=1}^{N}\left\{\rho_k\frac{L_{\text{pack}}}{C}+\frac{L_{\text{m}}}{C}\right\}=\rho_j.\qquad(6)$$

Multiplying Eq. (5) by λ_j and Eq. (6) by λ_i and subtracting one equation from the other, we obtain $\rho_i\lambda_j=\rho_j\lambda_i$, therefore: $\rho_i=\rho_j\dfrac{\lambda_i}{\lambda_j}$.

Similarly, we can express all ρ_k, $k=\overline{1,N}$ via $\rho_1:\rho_k=\rho_1\dfrac{\lambda_k}{\lambda_1}$. Substituting these relations into the first equation of the system of Eqs. (4), we obtain:

$$\lambda_1\sum_{k=1}^{N}\left\{\rho_1\frac{\lambda_k}{\lambda_1}\frac{L_{\text{pack}}}{C}+\frac{L_{\text{m}}}{C}\right\}=\rho_1,$$
$$\rho_1\sum_{k=1}^{N}\left\{\lambda_k\frac{L_{\text{pack}}}{C}\right\}+\frac{NL_{\text{m}}}{C}\lambda_1=\rho_1,\qquad(7)$$

$$\rho_1=\frac{N_{\text{node}}L_{\text{m}}}{C\left(1-\sum\limits_{k=1}^{N}\left\{\lambda_k\dfrac{L_{\text{pack}}}{C}\right\}\right)}\lambda_1,\qquad(8)$$

$$Z^{(1)}=\frac{\rho_1}{\lambda_1}=\frac{N_{\text{node}}L_{\text{m}}}{C\left(1-\sum\limits_{k=1}^{N}\left\{\lambda_k\dfrac{L_{\text{pack}}}{C}\right\}\right)}.\qquad(9)$$

The loading of the transmission medium is equal to: $R_{\text{medium}} = \sum_{k=1}^{N} \lambda_k \frac{L_{\text{pack}}}{C}$.

In steady-state mode, the load of the medium is equal to the probability that packets are transmitted in the medium. Accordingly, the value $1 - \sum_{k=1}^{N} \lambda_k \frac{L_{\text{pack}}}{C}$ is the probability that the medium is free from packet transmission.

As can be seen, Eq. (9) coincides completely with Eq. (1), which proves the first part of the token network polling cycle theorem.

Functional equations for determining the cycle of local computing systems with token access method taking into account emerging failures have the form [1]:

$$\begin{cases} Z^*_{\text{fail}}(s) = Z^*(\lambda_{\text{fail}} + s - \lambda_{\text{fail}}\Gamma^*_{\text{fail}}(s)), \\ \Gamma^*_{\text{fail}}(s) = F^*_{\text{fail}}(\lambda_{\text{fail}} + s - \lambda_{\text{fail}}\Gamma^*_{\text{fail}}(s)), \end{cases}$$

$$\begin{cases} Z^*_{\text{fail}}(s) = \int_0^\infty e^{-st} dZ_{\text{fail}}(t), \\ \Gamma^*_{\text{fail}}(s) = \int_0^\infty e^{-st} d\Gamma_{\text{fail}}(t), \end{cases} \quad (10)$$

$$\begin{cases} Z^*(\lambda_{\text{fail}} + s - \lambda_{\text{fail}}\Gamma^*_{\text{fail}}(s)) = \int_0^\infty e^{-(\lambda_{\text{fail}}+s-\lambda_{\text{fail}}\Gamma^*_{\text{fail}}(s))t} dZ(t), \\ F^*_{\text{fail}}(\lambda_{\text{fail}} + s - \lambda_{\text{fail}}\Gamma^*_{\text{fail}}(s)) = \int_0^\infty e^{-(\lambda_{\text{fail}}+s-\lambda_{\text{fail}}\Gamma^*_{\text{fail}}(s))t} dF(t), \end{cases}$$

where $Z_{\text{fail}}(t)$ is the distribution function (DF) of the local area network cycle taking into account failures; $Z(t)$ is the DF of the local area network cycle under conditions of reliable operation; $F_{\text{fail}}(t)$ is the DF of the transmission medium recovery time after failures; $\Gamma_{\text{fail}}(t)$ is the DF of the transmission medium occupancy period after failures; s is the complex parameter of the Laplace transform.

Let us differentiate the first functional Eq. (10) by s:

$$\left(Z^*_{\text{fail}}(s)\right)' = \left(\int_0^\infty e^{-(\lambda_{\text{fail}}+s-\lambda_{\text{fail}}G^*_{\text{fail}}(s))t} dZ(t)\right)' =$$

$$= -\int_0^\infty t e^{-(\lambda_{\text{fail}}+s-\lambda_{\text{fail}}G^*_{\text{fail}}(s))t} \left(1 - \lambda_{\text{fail}}\left(G^*_{\text{fail}}(s)\right)'\right) dZ(t),$$

at $s \to 0$ $(Z^*_{\text{fail}}(s))'|_{s=0} = -Z^{(1)}_{\text{fail}}$,

$\lambda_{\text{fail}} + s - \lambda_{\text{fail}}\Gamma^*_{\text{fail}}(s)|_{s=0} = 0$, $(\Gamma^*_{\text{fail}}(s))'\Big|_{s=0} = -\Gamma^{(1)}_{\text{fail}}$,

$$-\int_0^\infty t e^{\left\{-(s+\lambda_{\text{fail}}-\lambda_{\text{fail}}\Gamma^*_{\text{fail}}(s))t\right\}} (1 - \lambda_{\text{fail}}(\Gamma_{\text{fail}}(s)')dZ(t)|_{s=0} =$$

$$= -Z^{(1)}(1 + \lambda_{\text{fail}}\Gamma^{(1)}_{\text{fail}}).$$

Therefore,

$$Z^{(1)}_{\text{fail}} = Z^{(1)}(1 + \lambda_{\text{fail}}\Gamma^{(1)}_{\text{fail}}). \quad (11)$$

Differentiating the second functional Eq. (10) by s and finding the limit at $s \to 0$, we obtain the following relation for determining the average value of the occupancy period of the transmission medium after failure:

$$\Gamma^{(1)}_{\text{fail}} = \frac{F^{(1)}_{\text{fail}}}{1 - \lambda_{\text{fail}}F^{(1)}_{\text{fail}}}. \quad (12)$$

Substituting Eq. (12) into Eq. (11), we obtain:

$$Z^{(1)}_{\text{fail}} = \frac{Z^{(1)}}{1 - \lambda_{\text{fail}}F^{(1)}_{\text{fail}}}. \quad (13)$$

Eq. (13) completely coincides with Eq. (2), which proves the second part of the theorem on the cycle of polling of marker networks with regard to failures.

4. PACKET SERVICE CYCLE DEPENDENCE ON THE NUMBER OF NODES FOR DIFFERENT TYPES OF TRANSMITTED PACKETS

The present work investigates information transmission processes in multimode fiber optic transmission media using mathematical methods for calculating network structures with token access method developed earlier by the authors [1]. In the calculations, a four-mode fiber optic transmission medium with a bandwidth of 280 Mbps for each mode was analyzed. The length of the fiber channel can reach 100 km as well as in Fiber Distributed Data Interface (FDDI) networks with tokenized access. The modes were parameterized as follows:

- Length of packets transmitted using the first mode $L_{\text{pack }1} = 1024$ bit; packet transmission time constraint $T_{\text{dir }1} = 0.00025$ s; intensity of packets arriving in the network from each node $\lambda_{1i} = 2000$ pack/s.
- Length of packets transmitted using the second mode $L_{\text{pack }2} = 2048$ bit; packet transmission time constraint $T_{\text{dir }2} = 0.0005$ s; intensity of packets arriving in the network from each node $\lambda_{2i} = 1100$ pack/s.
- Length of packets transmitted using the third mode $L_{\text{pack }3} = 4096$ bit; packet transmission time constraint $T_{\text{dir }3} = 0.00075$ s; intensity of packets arriving in the network from each node $\lambda_{3i} = 590$ pack/s.
- Length of packets transmitted using the fourth mode $L_{\text{pack }4} = 8192$ bit; packet transmission time constraint $T_{\text{dir }4} = 0.001$ s; intensity of packets arriving in the network from each node $\lambda_{1i} = 300$ pack/s.

Table. Dependence of packet service time in multimode ring network on failure rate

λ_{fail}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	Without failure
T (L_{pack} = 1024 bit), s	0.003408	0.0004444	0.0001485	0.0001189	0.0001156
T (L_{pack} = 2048 bit), s	0.003765	0.000552	0.0002312	0.0001991	0.0001955
T (L_{pack} = 4096 bit), s	0.004485	0.0008598	0.0004978	0.0004617	0.0004576
T (L_{pack} = 8192 bit), s	0.004669	0.001029	0.0006856	0.0006513	0.0006475

A token-based access method was used to control the transmission of different classes of packets by means of a four-mode fiber ring. The length of the token L_m = 96 bit; the bandwidth of each mode is 280 Mbit/s. In the calculations, it was assumed that the MTBF of the transmission medium can take values of 10^4 s, 10^5 s, 10^6 s, 10^7 s and ∞; the failure recovery time is T_{rec} = 60 s.

In the calculations, the dependence of the packet service cycle, node and transmission medium load, packet service time, the probability of timely service of packets transmitted using different modes, and the performance of a multimode fiber optic network on the number of nodes connected to the transmission medium was investigated. The corresponding graphs are shown in Figs. 1–6.

The dependence of packet service time in a multimode ring network in the critical area of operation (at N_{node} = 110) on the intensity of failures is presented in Table.

As can be seen from the graphs (Figs. 1–6), as the number of nodes increases (with network scaling) in a multimode fiber optic network using the token access method, the load of the transmission medium and network nodes, the node polling cycle, and the transmission time of packets transmitted by different modes also increase; after reaching its maximum value, the network performance begins to drop sharply.

With increasing failure rates in the critical region of operation, with high transmission medium and node loads, all network characteristics change dramatically, in particular, packet service time.

Note that the maximum performance for all four types of transmitted packet is achieved when the number of nodes N_{node} = 100. The amount of timely transmitted information V_{pack} is maximized when the packet length L_{pack} = 8192:

$$V_{pack} = \lambda_{tot} L_{pack}.$$

The packet service time, which increases with packet size, is the longest at a high failure rate of 10^{-4}. For a packet of 1024 bits, the average service time is 0.003408 s, while for a packet of 8192 bits, it is already 0.004669 s.

When the failure rate is reduced to 10^{-5} the packet service time is significantly reduced. For a 1024-bit packet it decreases by almost 8 times (to 0.0004444 s), while for an 8192-bit packet, it decreases more than 4 times (to 0.001029 s).

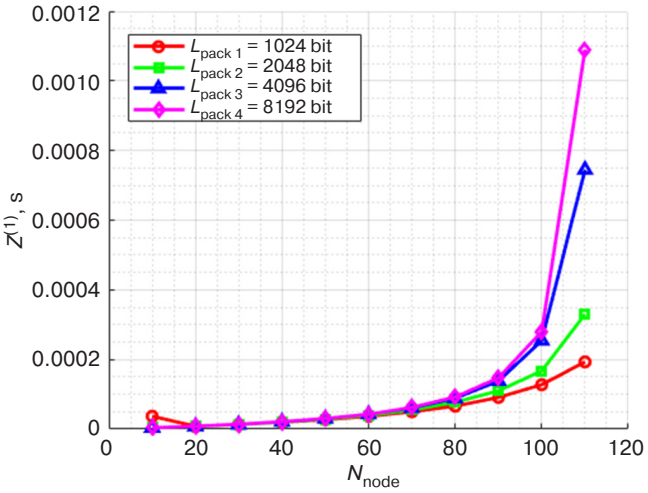


Fig. 1. Dependence of the service cycle $Z^{(1)}$ on the number of network nodes N_{node}

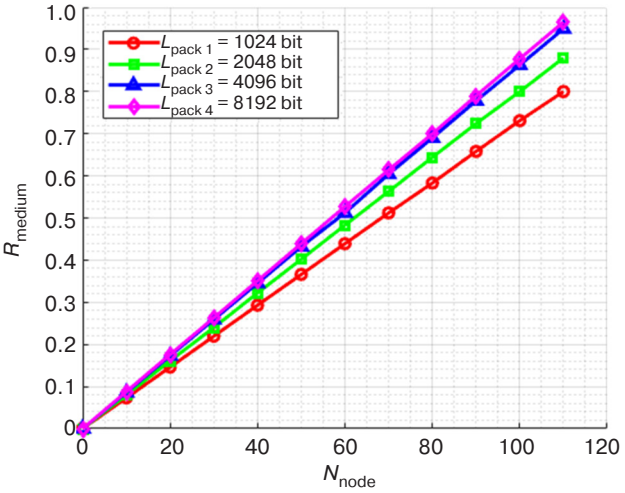


Fig. 2. Dependence of the transmission medium load R_{medium} for each mode on the number of nodes N_{node}

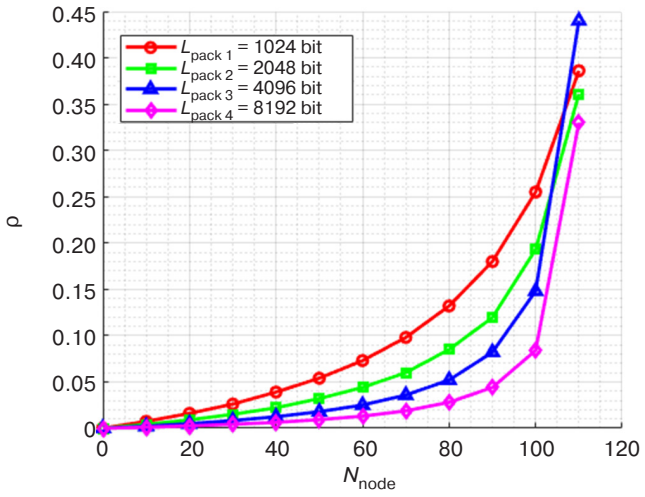


Fig. 3. Dependence of node loading for each mode ρ on the total number of nodes N_{node}

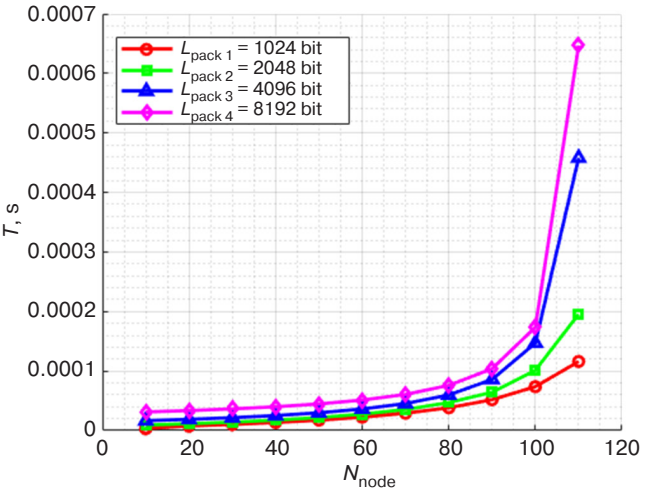


Fig. 4. Dependence of average service time of packets of 4 types *T* in multimode fiber optic network with token access method on the number of nodes N_{node}

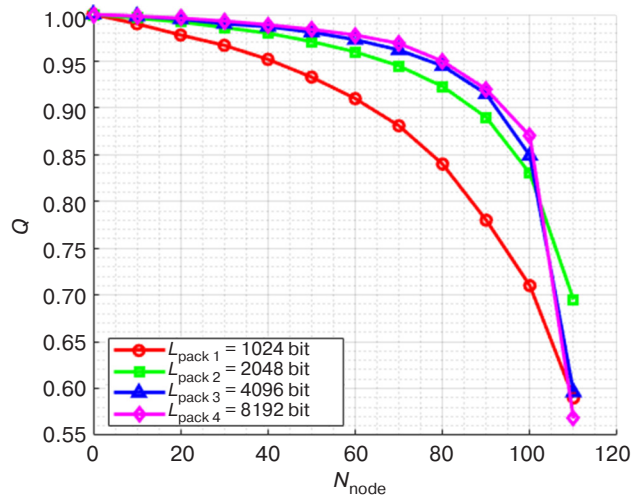


Fig. 5. Dependence of the probability of timely packet transmission *Q* in a multimode ring network on the number of nodes N_{node}

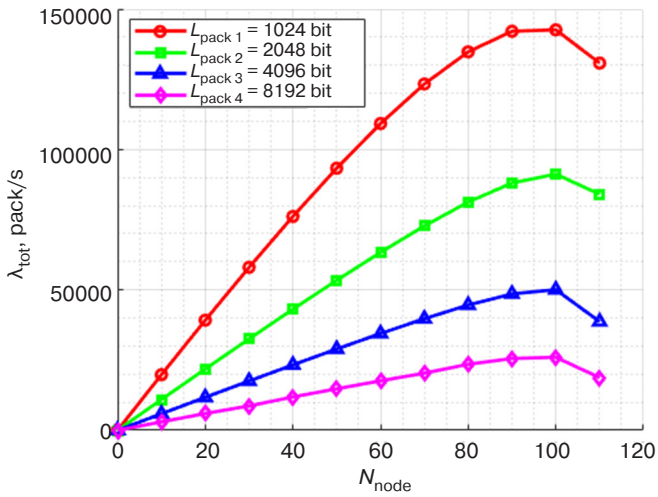


Fig. 6. Dependence of the performance λ_{tot} for each mode on the number of network nodes N_{node}

At failure rates of 10^{-6} and 10^{-7}, the processing time continues to decrease, but not as dramatically as when going from 10^{-4} to 10^{-5}. For example, for a 1024-bit packet, the time drops from 0.0001485 s to 0.0001189 s. This is already a less pronounced decrease.

In the case of no failures, the service time is stable and significantly lower than in the presence of failures. For example, for all packets from 1024 to 8192 bits, service time varies in the range of 0.0001156–0.0006475 s.

Failure intensity, which significantly affects the packet processing time in a multimode ring network, increases significantly at high intensities, especially for large packets. However, when the failure rate decreases to 10^{-6} and below, the processing time approaches the values characteristic of a network without failures.

CONCLUSIONS

This study highlights the key role of development and analysis of a mathematical model for optimization of data transmission processes in ring networks based on multimode fiber optic technology using the token access method.

The use of reliability-theory-, random-process- and mass-service methods, as well as the Laplace–Stieltjes transform, provided a basis for accurate mathematical equations that provide a deep analysis of the network characteristics to identify the key regularities of its operation.

The developed mathematical model takes into account critical parameters of multimode fiber optic medium including bandwidth, time delays, and data loss rate.

In the study, a comprehensive analysis of the operation of a multimode fiber optic network with token access method was carried out, in which the probabilistic and temporal characteristics of information transmission

for different classes of packets were studied. In particular, the dependencies of packet service time on the number of nodes connected to the transmission medium, as well as the influence of the bandwidth of each mode and the intensity of packet arrival in the network were considered. The analysis showed that as the number of nodes in the network increases, the utilization of the transmission medium and nodes increases, which, in turn, increases the service cycle and packet transmission time, and decreases the network performance. In addition, increasing the intensity of failures in the critical area of operation dramatically deteriorates the performance of the network – in particular, in terms of increased service time of packets and consequently reduced probability of their timely transmission.

We pay special attention to the issues of network reliability, namely the impact of failures of individual elements on its performance. This aspect is of particular importance for critical applications, where the reliability of data transmission determines the success of the entire system. By introducing reliability characteristics into the model, we were able to develop mechanisms to account for failures and propose ways to minimize them, which increases the resilience of the network to external influences and potential failures.

The results show that by properly tuning the parameters it is possible to significantly improve the overall network performance, reduce the probability of collisions, and increase the throughput. These conclusions confirm the high efficiency of the proposed solutions and their importance for improving network operation quality.

Based on the calculations and simulations, multimode fiber optic medium is confirmed as a promising solution for organizing high-performance local area networks, especially under conditions of increased reliability and information-loss minimization requirements. The developed mathematical model can serve as a universal tool for the analysis, construction and optimization of networks, as well as providing science-based recommendations for their effective use.

Authors' contributions

D.V. Zhmatov has developed the theoretical framework for the study by proving a key theorem on polling cycles in token-passing networks (Eqs. 1–13), calculated the network timing characteristics analytically and plotted the resulting dependencies (Figs. 1–6), analyzed the optimization results for the token-based access method and proposed recommendations to enhance network efficiency.

A.S. Leontiev has defined the problem statement, including data transmission requirements for networks, performed numerical simulations of network parameters (node load, throughput, and probability of on-time delivery), assessed the impact of transmission medium failures on network performance (Table).

REFERENCES

1. Leontyev A.S., Zhmatov D.V. Analytical method for analyzing message transmission processes in FDDI networks for digital substations. *Russ. Technol. J.* 2024;12(6):26–38. https://doi.org/10.32362/2500-316X-2024-12-6-26-38
2. Leontyev A.S. Development of Analytical Methods, Models, and Techniques for Local Area Networks Analysis. In: *Theoretical Issues of Software Engineering: Interuniversity Collection of Scientific Papers*. Moscow: MIREA; 2001. P. 70–94 (in Russ.).
3. Leontyev A.S. Multilevel Analytical and Analytical-Simulation Models for Evaluating the Probabilistic and Temporal Characteristics of Multimachine Computing Complexes with Regard to Reliability. *Mezhdunarodnyi nauchno-issledovatel'skii zhurnal = International Research Journal*. 2023;5(131) (in Russ.). https://doi.org/10.23670/IRJ.2023.131.8
4. Sapna, Sharma M. Performance evaluation of a wired network with & without Load Balancer and Firewall. In: *2010 International Conference on Electronics and Information Engineering.* Kyoto, Japan. 2010. P. V2-515–V2-519. https://doi.org/10.1109/ICEIE.2010.5559755
5. Wei M., Chen Z. Study of LANs access technologies in wind power system. In: *IEEE PES General Meeting*. Minneapolis, MN, USA. 2010.
6. Zvonareva G.A., Buzunov D.S. Using Simulation Modeling to Estimate Time Characteristics of a Distributed Computing System. *Otkrytoe obrazovanie = Open Education*. 2022;26(5):32–39 (in Russ.). https://doi.org/10.21686/1818-4243-2022-5-32-39
7. Kul'ba V.V., Mamikonov A.G., Shelkov A.B. Redundancy of program modules and data arrays in a MIS. *Avtomatika i telemekhanika = Automation and Remote Control*. 1980;8:133–141 (in Russ.).
8. Talalaev A.A., Frolenko V.P. Fault-tolerant system for organizing high-performance computing for solving data stream processing problems. *Programmnye sistemy: Teoriya i prilozheniya = Program Systems: Theory and Applications*. 2018;9(1–36):85–108 (in Russ.). https://doi.org/10.25209/2079-3316-2018-9-1-85-108
9. Akimova G.P., Solovyev A.V., Tarkhanov I.A. Modeling the reliability of distributed information systems. *Informatsionnye tekhnologii i vychislitel'nye sistemy = Journal of Information Technologies and Computing Systems*. 2019;3:70–86 (in Russ.). https://doi.org/10.14357/20718632190307

10. Pavsky V.A., Pavsky K.V. Mathematical Model for Calculating Reliability Indicators of Scalable Computer Systems Considering Switching Time. *Izvestiya YuFU. Tekhnicheskie nauki = Izvestiya SFedU. Engineering Sciences*. 2020;2(212): 134–145 (in Russ.). https://doi.org/10.18522/2311-3103-2020-2-134-145
11. Waseem A., Wu Y.W. A survey on reliability in distributed systems. *J. Comput. Syst. Sci*. 2013;79(8):1243–1255. https://doi.org/10.1016/j.jcss.2013.02.006
12. Ivanichkina L.V., Neporada A.L. The Reliability model of a distributed data storage in case of explicit and latent disk faults. *Trudy Instituta sistemnogo programmirovaniya RAN = Proceedings of the Institute for System Programming of the RAS*. 2015;27(6):253–274 (in Russ.). https://doi.org/10.15514/ISPRAS-2015-27(6)-16
13. Rothe S., Besser K.L., Krause D., Kuschmierz R., Koukourakis N., Jorswieck E., Czarske J.W. Securing Data in Multimode Fibers by Exploiting Mode-Dependent Light Propagation Effects. *Research*. 2023;6:Article 0065. https://doi.org/10.34133/research.0065
14. Ferguson K., Kleinrock L. Optimal update timing of stale information metrics, including age of information. *IEEE Journal on Selected Areas in Information Theory.* 2023;4:734–746. https://doi.org/10.1109/JSAIT.2023.3344760
15. He Z., Kleinrock L. Optimization of Assisted Search Over Server-Mediated Peer-to-peer Networks. In: *2022 IEEE Global Communications Conference* (*GLOBECOM 2022*). IEEE. 2022. P. 4928–4934. https://doi.org/10.1109/GLOBECOM48099.2022.10000846
16. Jain R. Error characteristics of fiber distributed data interface (FDDI). *IEEE Trans. Commun*. 1990;38(8):1244–1252. https://doi.org/10.1109/26.58757
17. Kleinrock L. Internet congestion control using power metric: *Keep the just full, but no fuller*. *Ad Hoc Networks.* 2018;80: 142–157. https://doi.org/10.1016/j.adhoc.2018.05.015

СПИСОК ЛИТЕРАТУРЫ

1. Леонтьев А.С., Жматов Д.В. Аналитический метод анализа процессов передачи сообщений в оптоволоконных сетях с маркерным доступом для цифровых подстанций. *Russ. Technol. J*. 2024;12(6):26–38. https://doi.org/10.32362/2500-316X-2024-12-6-26-38
2. Леонтьев А.С. Разработка аналитических методов, моделей и методик анализа локальных вычислительных сетей. *Теоретические вопросы программного обеспечения: Межвузовский сборник научных трудов*. М.: МИРЭА; 2001. С. 70–94.
3. Леонтьев А.С. Многоуровневые аналитические и аналитико-имитационные модели оценки вероятностно-временных характеристик многомашинных вычислительных комплексов с учетом надежности. *Международный научно-исследовательский журнал.* 2023;5(131). https://doi.org/10.23670/IRJ.2023.131.8
4. Sapna, Sharma M. Performance evaluation of a wired network with & without Load Balancer and Firewall. In: *2010 International Conference on Electronics and Information Engineering.* Kyoto, Japan. 2010. P. V2-515–V2-519. https://doi.org/10.1109/ICEIE.2010.5559755
5. Wei M., Chen Z. Study of LANs access technologies in wind power system. In: *IEEE PES General Meeting*. Minneapolis, MN, USA. 2010.
6. Звонарева Г.А., Бузунов Д.С. Использование имитационного моделирования для оценки временных характеристик распределенной вычислительной системы. *Открытое образование*. 2022;26(5):32–39. https://doi.org/10.21686/1818-4243-2022-5-32-39
7. Кульба В.В., Мамиконов А.Г., Шелков А.Б. Резервирование программных модулей и информационных массивов в АСУ. *Автоматика и телемеханика*. 1980;8:133–141.
8. Талалаев А.А., Фроленко В.П. Отказоустойчивая система организации высокопроизводительных вычислений для решения задач обработки потоков данных. *Программные системы: Теория и приложения*. 2018;9(1–36):85–108. https://doi.org/10.25209/2079-3316-2018-9-1-85-108
9. Акимова Г.П., Соловьев А.В., Тарханов И.А. Моделирование надежности распределенных вычислительных систем. *Информационные технологии и вычислительные системы* (*ИТиВС*). 2019;3:70–86. https://doi.org/10.14357/20718632190307
10. Павский В.А., Павский К.В. Математическая модель для расчета показателей надежности масштабируемых вычислительных систем с учетом времени переключения. *Известия ЮФУ. Технические науки*. 2020;2(212):134–145. https://doi.org/10.18522/2311-3103-2020-2-134-145
11. Waseem A., Wu Y.W. A survey on reliability in distributed systems. *J. Comput. Syst. Sci*. 2013;79(8):1243–1255. https://doi.org/10.1016/j.jcss.2013.02.006
12. Иваничкина Л.В., Непорада А.Л. Модель надежности распределенной системы хранения данных в условиях явных и скрытых дисковых сбоев. *Труды Института системного программирования РАН*. 2015;27(6):253–274. https://doi.org/10.15514/ISPRAS-2015-27(6)-16
13. Rothe S., Besser K.L., Krause D., Kuschmierz R., Koukourakis N., Jorswieck E., Czarske J.W. Securing Data in Multimode Fibers by Exploiting Mode-Dependent Light Propagation Effects. *Research*. 2023;6:Article 0065. https://doi.org/10.34133/research.0065
14. Ferguson K., Kleinrock L. Optimal update timing of stale information metrics, including age of information. *IEEE Journal on Selected Areas in Information Theory.* 2023;4:734–746. https://doi.org/10.1109/JSAIT.2023.3344760

15. He Z., Kleinrock L. Optimization of Assisted Search Over Server-Mediated Peer-to-peer Networks. In: *2022 IEEE Global Communications Conference* (*GLOBECOM 2022*). IEEE. 2022. P. 4928–4934. https://doi.org/10.1109/GLOBECOM48099.2022.10000846
16. Jain R. Error characteristics of fiber distributed data interface (FDDI). *IEEE Trans. Commun*. 1990;38(8):1244–1252. https://doi.org/10.1109/26.58757
17. Kleinrock L. Internet congestion control using power metric: *Keep the just full, but no fuller*. *Ad Hoc Networks.* 2018;80:142–157. https://doi.org/10.1016/j.adhoc.2018.05.015

About the Authors

Dmitry V. Zhmatov, Cand. Sci. (Eng.), Associate Professor, Department of Mathematical Support and Standardization, Institute of Information Technologies, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: zhmatov@mirea.ru. Scopus Author ID 56825948100, RSCI SPIN-code 2641-6783, https://orcid.org/0000-0002-7192-2446

Alexander S. Leontyev, Cand. Sci. (Eng.), Senior Researcher, Associate Professor, Department of Mathematical Support and Standardization, Institute of Information Technologies, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: leontev@mirea.ru. RSCI SPIN-code 5798-9721, https://orcid.org/0000-0003-3673-2468

Об авторах

Жматов Дмитрий Владимирович, к.т.н., доцент, доцент кафедры математического обеспечения и стандартизации информационных технологий, Институт информационных технологий, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: zhmatov@mirea.ru. Scopus Author ID 56825948100, SPIN-код РИНЦ 2641-6783, https://orcid.org/0000-0002-7192-2446

Леонтьев Александр Савельевич, к.т.н., старший научный сотрудник, доцент кафедры математического обеспечения и стандартизации информационных технологий, Институт информационных технологий, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: leontev@mirea.ru. SPIN-код РИНЦ 5798-9721, https://orcid.org/0000-0003-3673-2468

Translated from Russian into English by L. Bychkova
Edited for English language and spelling by Thomas A. Beavitt

Modern radio engineering and telecommunication systems

Современные радиотехнические и телекоммуникационные системы

UDC 621.396.65
https://doi.org/10.32362/2500-316X-2025-13-5-51-62
EDN NOGIBI

RESEARCH ARTICLE

Analysis of time software and hardware delays in audio module circuits with cyber-physical SPICE emulation

Nikita R. Levchenko @, Mihail S. Kostin

MIREA – Russian Technological University, Moscow, 119454 Russia
@ Corresponding author, e-mail: Levchenko_n@mirea.ru

• **Submitted:** 04.03.2025 • **Revised:** 05.05.2025 • **Accepted:** 26.07.2025

Abstract

Objectives. The study sets out to parametrically investigate the impact of time delays within cyber-physical emulation circuits for signal audio modules. Specifically, it examines how delays introduced by analog-to-digital and digital-to-analog converters of the hardware/software interface, the central processor, and the visual graphical emulation (VGE) software environment are influenced by factors like the selected data input-output protocol and the VGE block preset configurations such as sampling rate, buffer size and time, and the number of channels.

Methods. Used methods of architectural SPICE[1] modeling of electrical circuits on the VGE *Simulink* software platforms leverage the resources of the Simscape library and *LiveSPICE*. Additional methods include those for incorporating differential equations in the numerical analysis of SPICE models designed for analog circuits and techniques for processing experimental data generated from cyber-physical emulation using the built-in *Simulink* environment and associated laboratory radio measurement tools.

Results. The study introduces a novel approach to emulate analog audio devices using cyber-physical SPICE modeling. Through the use of digital twins, the study investigates the impact of modifiable parameters on signal delays within audio module circuits during cyber-physical emulation. Based on these findings, technical guidelines are provided for selecting appropriate delay correction settings between 20 and 120 ms to ensure efficient high-speed audio signal post-processing.

Conclusions. By configuring the VGE software block's settings identically to the ASIO[2] data input/output protocol prevalent in audio interface technology (44.1 kHz sampling frequency, 8 buffer size) substantially decreased latency in typical audio module circuit nodes is achieved with cyber-physical emulation built into the VGE *LiveSPICE* environment. The achieved time delays of 5 ms direct transmission circuit contrast with the 7 ms latency observed in the cyber-physical emulation of the SPICE circuit when both are benchmarked within the VGE *Simulink* environment. The

[1] SPICE (Simulation Program with Integrated Circuit Emphasis) is an open-source simulator for general-purpose electronic circuits.

[2] Audio Stream Input/Output is a low-latency data transmission technology providing applications with a uniform interface to hardware resources.

successful implementation of cyber-physical emulation for SPICE models is achieved through the use of particular settings, such as a 44.1 kHz sampling frequency, buffer sizes ranging from 512 to 1024 samples, and the use of the ASIO data input/output protocol.

Keywords: time delay, correction, cyber-physical SPICE emulation, Simulink, LiveSPICE, Simscape, audio interface, ASIO, signal audio module

For citation: Levchenko N.R., Kostin M.S. Analysis of time software and hardware delays in audio module circuits with cyber-physical SPICE emulation. *Russian Technological Journal.* 2025;13(5):51–62. https://doi.org/10.32362/2500-316X-2025-13-5-51-62, https://www.elibrary.ru/NOGIBI

Financial disclosure: The authors have no financial or proprietary interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Анализ временных программно-аппаратных задержек в схемах аудиомодулей с киберфизической SPICE-эмуляцией

Н.Р. Левченко [@], М.С. Костин

МИРЭА – Российский технологический университет, Москва, 119454 Россия
[@] Автор для переписки, e-mail: Levchenko_n@mirea.ru

• **Поступила:** 04.03.2025 • **Доработана:** 05.05.2025 • **Принята к опубликованию:** 26.07.2025

Резюме

Цели. Цель статьи – параметрический анализ влияния временных задержек в схемах киберфизической эмуляции сигнальных аудиомодулей, вносимых аналого-цифровыми и цифро-аналоговыми преобразователями программно-аппаратного интерфейса, центральным процессором и программной средой визуально-графической эмуляции (ВГЭ) в зависимости от выбранного протокола «ввода-вывода» данных и установленных преднастроек программного блока ВГЭ, таких как частота дискретизации, размер и время буфера, число каналов.

Методы. Применяются методы архитектурного SPICE[3]-моделирования электрических схем на программных платформах ВГЭ *Simulink*, в т.ч. с использованием ресурсов библиотеки Simscape, и *LiveSPICE*; методы интегрирования дифференциальных уравнений при численном анализе SPICE-моделей аналоговых схем; методы обработки экспериментальных данных киберфизической эмуляции с помощью встроенных средств среды *Simulink* и лабораторного радиоизмерительного оборудования.

Результаты. Предложен метод киберфизической SPICE-эмуляции аналоговых аудиоустройств. Получены результаты анализа формирования временных задержек в схемах сигнальных аудиомодулей с киберфизической эмуляцией при вариации преднастраиваемых параметров, влияющих на задержки сигналов, с применением двойников. Разработаны технические рекомендации выбора корректирующих параметров временных задержек от 20 до 120 мс для обеспечения постобработки аудиосигнала.

Выводы. Показано, что для часто используемого в аудиоинтерфейсной технике протокола «ввода-вывода» данных ASIO[4] при тождественно установленных преднастройках программного блока ВГЭ (частота

[3] SPICE (англ. Simulation Program with Integrated Circuit Emphasis) – программа-симулятор электронных схем общего назначения с открытым исходным кодом. [SPICE (Simulation Program with Integrated Circuit Emphasis) is an open source simulator of general-purpose electronic circuits.]

[4] Audio Stream Input/Output, «ввод-вывод потоковых аудиоданных» – протокол передачи данных с малой задержкой, предоставляющий приложениям унифицированный интерфейс к аппаратным ресурсам. [Audio Stream Input/Output is a low-latency data transmission technology providing applications with a uniform interface to hardware resources.]

дискретизации 44.1 кГц, размер буфера 8) типовые функциональные узлы схем аудиомодулей с киберфизической эмуляцией, построенные в среде ВГЭ *LiveSPICE*, имеют наименьшие временные задержки 5 мс – для схемы прямого прохождения сигнала и 7 мс – в случае с киберфизической эмуляцией SPICE-схемы в отношении к их реализации в среде ВГЭ *Simulink*. Установлено, что обоснованно выбранными настройками при практической реализации метода киберфизической эмуляции SPICE-моделей являются: частота дискретизации 44.1 кГц, размер буфера от 512 до 1024 семплов и протокол «ввода-вывода» данных ASIO.

Ключевые слова: временные задержки, киберфизическая SPICE-эмуляция, Simulink, LiveSPICE, Simscape, аудиоинтерфейс, ASIO, сигнальный аудиомодуль

Для цитирования: Левченко Н.Р., Костин М.С. Анализ временных программно-аппаратных задержек в схемах аудиомодулей с киберфизической SPICE-эмуляцией. *Russian Technological Journal*. 2025;13(5):51–62. https://doi.org/10.32362/2500-316X-2025-13-5-51-62, https://www.elibrary.ru/NOGIBI

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

A widely recognized classification for digital representations of electrical circuits used in automated circuit design and software-based numerical analysis of functional units within radio-electronic equipment includes code-based, visual-graphical, and hybrid approaches [1]. For example, the use of program code to interpret the hardware architecture of functional circuits in signal radio modules necessitates the utilization of specialized languages like SPICE[5], VHDL[6], or object-oriented programming languages such as C/C++, Java, Python, Ruby, etc. Visual-graphical descriptions offer code-based interpretations of circuit and functional architectures by employing functional blocks connected in a directed graph. Mixed emulation involves a variable combination of code, block, and specialized hardware description languages to represent circuit elements along with their corresponding electrical and logical interconnections.

The SPICE emulator [2], representing a typical means of interpreting analog circuit solutions, is based on numerical methods for integrating differential systems of equations in order to determine the currents and voltages in a given electrical circuit.

Electronic component base (ECB) manufacturers, including Texas Instruments[7], Linear Technology[8], KYOCERA AVX[9], STMicroelectronics[10] and others, offer design engineers access to parametric SPICE models and vector S-parameter models for ECB libraries, which are all backed by experimentally validated performance data. By leveraging original ECB libraries with comprehensive parameterization tailored to specific equivalent models, which are constructed upon a popular SPICE emulator platform, an electronic printed circuit board's radio-technical attributes can be closely aligned with the outcomes of its circuit emulation within a condensed timeframe.

Examples of visual-graphical design tools include *SigmaStudio*[11], *LTSPICE*[12], *QSPICE*[13], *Tina-TI*[14], and *NI Multisim*[15]. These automated visual design platforms, which operate with hidden code, facilitate the creation of electronic product circuit architectures of any configuration by utilizing library blocks and modules. However, a drawback of these automated platforms is their inability to enable real-time modifications to block descriptions through user-defined program code. Platforms like *LabView*[16], *Simulink*, *Proteus Design Suite*[17], and some others circumvent this limitation with visual graphical emulation (VGE). VGE platforms, which may be categorized as mixed description, allow

[5] SPICE (Simulation Program with Integrated Circuit Emphasis) is an open source simulator for general-purpose electronic circuits.

[6] VHDL is a Very High Speed Integrated Circuit (VHSIC) Hardware Description Language.

[7] Texas Instruments official website. https://www.ti.com/. Accessed February 20, 2025.

[8] https://lineartech.com/. Accessed February 20, 2025.

[9] https://www.kyocera-avx.com/. Accessed February 20, 2025.

[10] https://www.st.com/. Accessed February 20, 2025.

[11] https://wiki.analog.com/resources/tools-software/sigmastudio. Accessed February 20, 2025.

[12] https://www.quadcept.com/en/service/simulation/ltspice/. Accessed February 20, 2025.

[13] https://www.qorvo.com/design-hub/design-tools/interactive/qspice. Accessed February 20, 2025.

[14] https://designsoftware.com/home/English/. Accessed February 20, 2025.

[15] https://www.ni.com/en/support/downloads/software-products/download.multisim.html?srsltid=AfmBOopwVYxFcVbXG5r0bM3lm203f24w4RwAIITsF3NAp9On-4a8eb_F#452133. Accessed February 20, 2025.

[16] https://www.ni.com/en/shop/labview.html?srsltid=AfmBOoowwWPAjm-J6TthCp-LDLqR2ikpuO0W5eJHBp2ukB0E3Xe00cfO. Accessed February 20, 2025.

[17] https://proteus.no/proteus.html. Accessed February 20, 2025.

modifications to blocks through code written in high-level or command-based programming languages.

The VGE *LiveSPICE*[18] environment, a cyber-physical platform developed by inventor Dillon Sharlet, provides a unique solution for emulating audio modules using SPICE. This software environment empowers users to design, analyze, and debug electrical audio circuits built on the ASIO[19] protocol. It facilitates this process by enabling real-time signal transmission between the audio interface and the SPICE model of the circuit.

Sharlet's cyber-physical SPICE emulation technique, which was initially applied within the VGE *LiveSPICE* platform, holds significant value for various research endeavors. By integrating physically authentic signals from external circuit-hardware solutions in real-time mode, its implementation can empower the development, debugging, and examination of analog low-frequency SPICE circuits. Nevertheless, implementing this method in an engineering setting necessitates a shift towards multifunctional VGE platforms that incorporate diverse circuit descriptions and a broader library selection. Furthermore, open access to user-friendly software, like *Simulink*, is crucial for this adaptation.

This technique substitutes the hardware module or functional component within a real radio-electronic device's circuit design with a cyber-circuit constructed using the VGE software platform, which is designed for signal analysis and debugging.

1. ANALYSIS OF THE IMPACT OF TEMPORARY DELAYS IN AUDIO SIGNAL DEVICE CIRCUITS WITH CYBER-PHYSICAL EMULATION

The functional circuit of the cyber-physical SPICE emulation method (Fig. 1) demonstrates signal transmission via the input-output (I/O) protocol connecting the software/hardware interface (audio interface) with the pre-installed VGE environment. The audio signal is directed to the input channels of the analog-to-digital converter (ADC) within the software/hardware interface. This interface, which communicates with a computer using a USB connection, sends signal data to the VGE system via the audio signal transmission protocol facilitated by a driver. The VGE system manipulates the signal and sends the modified signal back to the interface, which subsequently directs it to a digital-to-analog converter (DAC) through designated output channels.

Beyond ADC/DAC performance and I/O protocol latencies, several key VGE software block parameters influence the time delays encountered as an input signal traverses the software/hardware interface, driver, and VGE system. These parameters include sampling frequency, buffer size and duration, the quantity of hardware channels utilized, and others. The total accumulated time delay, denoted as $T_{\text{TD accum}}$, can be expressed using the following general formula:

$$T_{\text{TD accum}} = T_{\text{HTD}} + T_{\text{TD VGE}}, \quad (1)$$

where T_{HTD} is hardware time delays and $T_{\text{TD VGE}}$ is time delays introduced by the VGE block.

The hardware time delays during audio interface processing, T_{INTF}, are described by the following equation:

$$T_{\text{INTF}} = T_{\text{DAC}} + T_{\text{ADC}} + T_{\text{Buff.}} + T_{\text{Res.}} + T_{\text{SR}}, \quad (2)$$

where T_{DAC} is the time delays introduced by DAC output channels of the software/hardware interface; T_{ADC} is the time delays introduced by ADC input channels of the software/hardware interface; $T_{\text{Buff.}}$ is the time delays introduced by the buffer of the software/hardware interface; $T_{\text{Res.}}$ is the time delays defined by the ADC/DAC resolution of the software/hardware interface I/O channel; T_{SF} is time delays defined by the sampling frequency.

Audio equipment exhibits a noticeable time lag when the delay exceeds 20 ms; delays shorter than this threshold are generally inaudible. Within circuit analysis, delays up to 130 ms[20] are considered acceptable.

Within the VGE *Simulink* platform, a plugin is developed to simulate signal audio modules in a cyber-physical context by analyzing time delays within these emulated circuits. The replacement of actual audio module with the digital SPICE twin facilitates real-time development and debugging of audio modules using the *Simulink*'s Simscape library. This approach is further utilized to investigate key characteristics mentioned in the introduction, such as buffer size and sampling frequency[21] [3, 4]. The functional circuit of the plugin is shown in Fig. 2.

[18] The *LiveSPICE* developer official website. https://www.livespice.org/. Accessed February 20, 2025.

[19] Audio Stream Input/Output is a low-latency data transmission technology providing applications with a uniform interface to hardware resources.

[20] What is the latency in audio recording? https://taplic.com/audio-tips/latency-in-audio/. Accessed February 20, 2025.

[21] Levchenko N.R., Kostin M.S., Filatov S.V., Pechenkin S.M. *Audio Interface Switching Program with SPICE Technology*: Certificate of State Registration of Computer Program 2024685518 RF. Publ. 30.10.2024 (in Russ.).

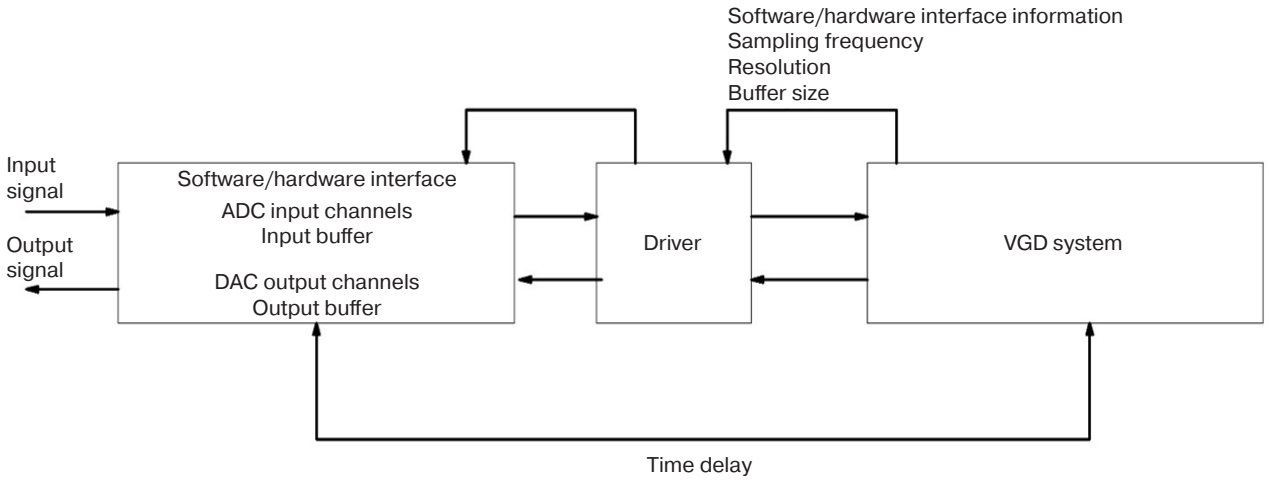


Fig. 1. Functional circuit for implementing the cyber-physical SPICE emulation method. The VGE system is a visual graphic design system

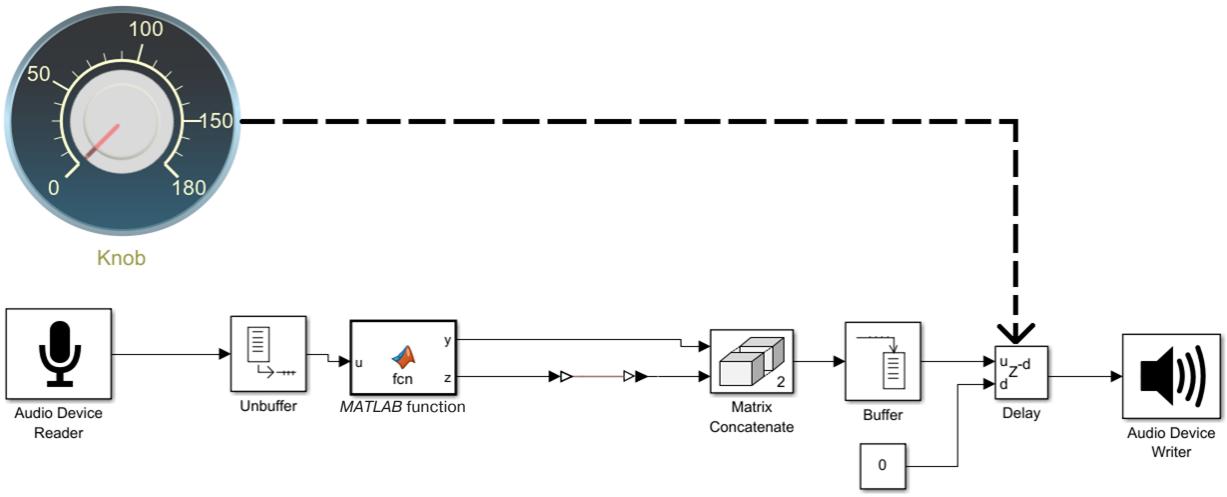


Fig. 2. Functional circuit of the plugin for cyber-physical emulation of signal audio modules in the VGE *Simulink* environment

This plugin comprises the following functional blocks:

1. "Audio Device Reader" block designed to initialize input channels for a software or hardware interface using a defined protocol.
2. "Unbuffer" block for transforming incoming frames from the software/hardware interface into a continuous stream of samples.
3. "*MATLAB* function" serving as a crucial block-level function, serving to isolate the channels within the software/hardware interface.
4. "Matrix Concatenate" block used to construct an $M \times N$ matrix, with M representing the number of samples and N denoting the quantity of channels utilized.
5. "Buffer" block used to convert the resulting matrix into frames, enabling subsequent transmission to the software or hardware interface.
6. "Delay" block used to counteract phase shift before the signal is sent to the software/hardware interface output channels.
7. "Audio Device Writer" block used for capturing incoming audio frames and sending them to the designated output channels of the software/hardware interface.
8. "Knob" block representing a digital potentiometer for adjusting the phase shift within the "Delay" block.

2. EXPERIMENTAL STUDY OF TIME DELAYS IN AUDIO MODULE CIRCUITS WITH CYBER-PHYSICAL SPICE EMULATION

The study evaluates the impact of hardware/software interface elements, including buffer size, resolution, sampling frequency, and audio data transmission

protocol, on the time delays experienced during signal transmission. Two distinct VGE platforms are evaluated: *LiveSPICE* and *Simulink*, together with its specialized plugin for cyber-physical SPICE emulation of signal audio modules. The selected software/hardware interface for this study is the UMC1820 audio interface (Behringer, Germany). Users have the flexibility to select from various parameters, including buffer size options ranging from 8 to 2048 samples, sampling frequencies at 44.1, 48.0, 88.2, and 96.0 kHz, and support for data transfer protocols like ASIO, WASAPI[22], and DirectSound [5–7].

The scheme of cyber-physical emulation of signal audio modules is shown in Fig. 3.

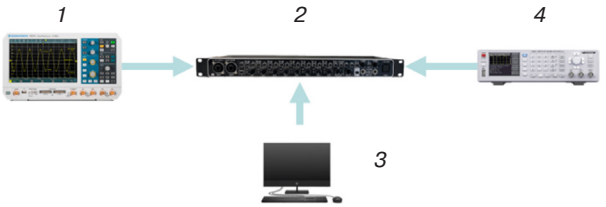


Fig. 3. Scheme of cyber-physical emulation used to explore time delays: *1* is oscilloscope RTB2002 (Rohde & Schwarz, Germany); *2* is audio interface UMC1820; *3* is personal computer; *4* is signal generator HMF2550 (Rohde & Schwarz, Germany)

The time delays τ_{DLY} obtained from a cyber-physical emulation setup are presented in Table 1. This setup employs varying sampling frequencies F_S while maintaining a constant buffer size of N = 2048 for a straightforward, single-channel audio stream directly from input to output of the audio interface.

Table 1. Outcomes of examining time delays that occur when adjusting sampling frequency while maintaining a constant buffer size

VGE environment, data I/O protocol							
LiveSPICE, ASIO		*Simulink*, ASIO		*Simulink*, DirectSound		*Simulink*, WASAPI	
F_S, kHz	τ_{DLY}, ms	F_S, kHz	τ_{DLY}, ms	F_S, kHz	τ_{DLY}, ms	F_S, kHz	τ_{DLY}, ms
44.1	81.8	44.1	128	44.1	174	44.1	135
48	78.6	48	121	48	162	48	130
88.2	71	88.2	104	88.2	110	88.2	108
96	70	96	98.8	96	108	96	100

The findings from research on the VGE *LiveSPICE* environment utilizing DirectSound and WASAPI protocols remain unpublished because the VGE *LiveSPICE* environment is inherently designed to operate exclusively with the ASIO protocol.

The time delays incurred by varying the buffer size N are presented in Table 2. These measurements are taken with a constant sampling frequency F_S = 44.1 kHz via a single-channel direct streaming method between the audio interface's input and output, which utilizes the ASIO, DirectSound, and WASAPI protocols.

Table 2. Outcomes of examining time delays that occur when adjusting the buffer size while maintaining a constant sampling frequency

VGE environment, data I/O protocol							
LiveSPICE, ASIO		*Simulink*, ASIO		*Simulink*, DirectSound		*Simulink*, WASAPI	
N	τ_{DLY}, ms	*N*	τ_{DLY}, ms	*N*	τ_{DLY}, ms	*N*	τ_{DLY}, ms
8	5	8	81.8	8	–	8	90
16	5	16	82.4	16	–	16	92
32	6	32	82.4	32	–	32	92
64	6.6	64	82.4	64	–	64	92
128	8.6	128	85.4	128	–	128	95
256	13.8	256	88.4	256	103	256	99
512	23.2	512	93.8	512	109	512	100
1024	42.8	1024	105.8	1024	111	1024	107
2048	82	2048	129	2048	164	2048	135

It can be seen from Tables 1 and 2 that adjusting the buffer size in the software/hardware interface for both the direct signal transmission method and the DirectSound audio protocol leads to longer transmission delays. Notably, the *Simulink* environment's direct signal pathway, when coupled with the cyber-physical emulation plugin and DirectSound transmission, exhibits the most significant delay, reaching 174 ms. Direct signal transmission within the *LiveSPICE* platform demonstrated very short delays down to 5 ms. Interestingly, buffer overflow leads to reduced time delays when the sampling frequency is elevated [8–10].

For assessing time delays within VGE systems using a digital SPICE twin, a standard low-frequency amplifier (LFA) circuit is utilized. The LFA fundamental circuit designs, which were developed on VGE *Simulink* platforms leveraging the Simscape library and *LiveSPICE*, are shown in Figs. 4 and 5. The use of a standard LFA circuit for analyzing and adjusting time delay properties in analog circuits with cyber-physical SPICE simulation is justified by the common circuit solutions of audio mixer consoles (such as active equalization lines, preamps, summers, panoramic dividers, and others) based on operational amplifiers (OAs) [11–13].

[22] Windows Audio Session API is a low-latency data transfer protocol providing applications with a unified interface to hardware resources, developed by Microsoft.

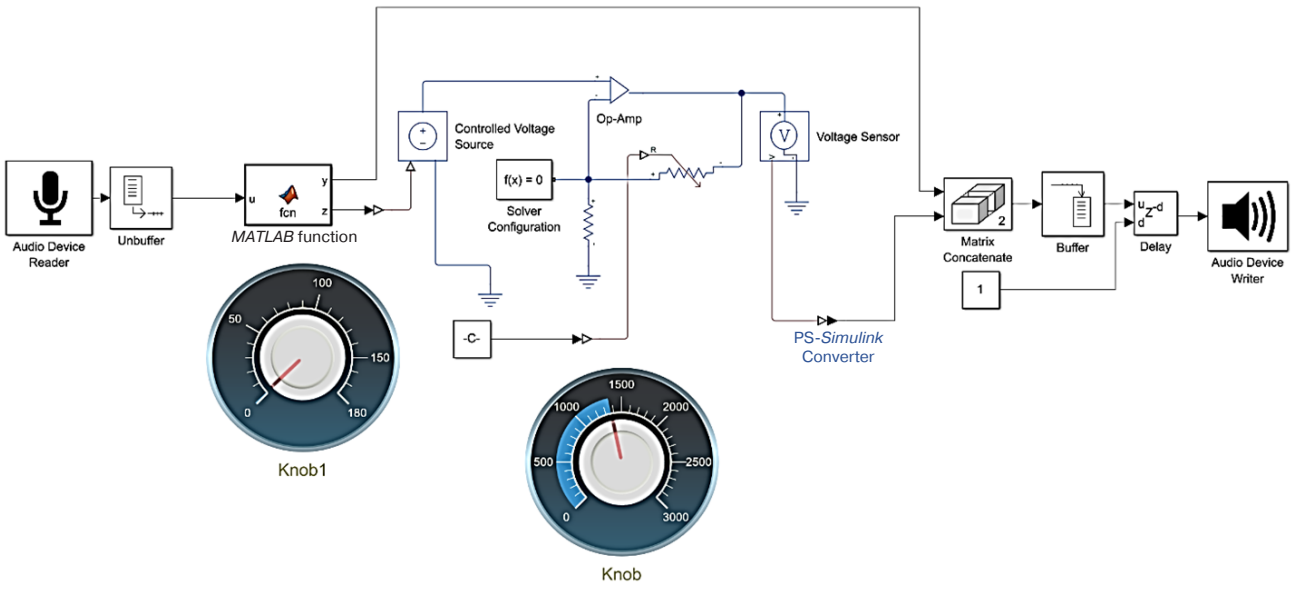


Fig. 4. LFA circuit in the VGE *Simulink* environment for analyzing time delays

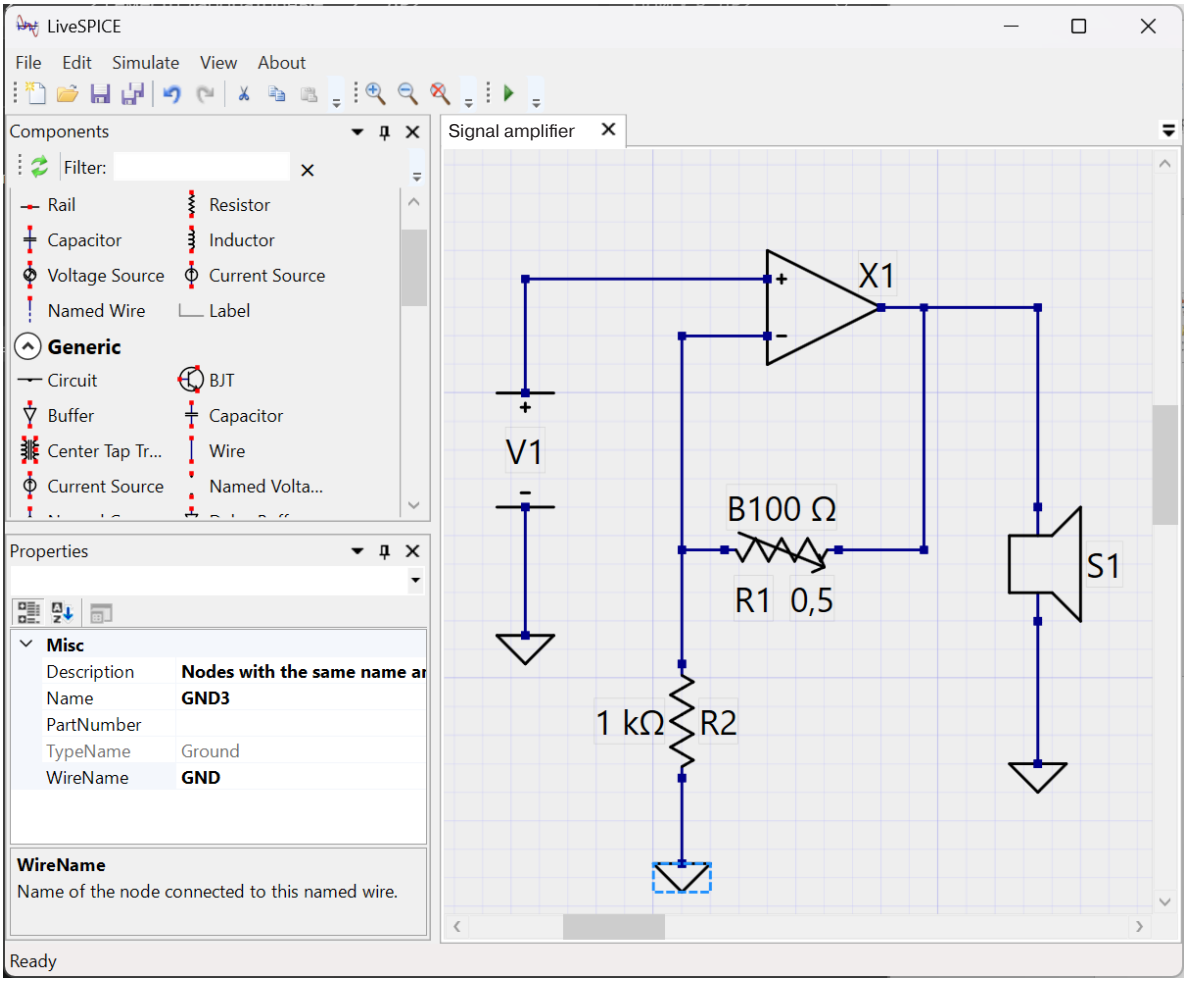


Fig. 5. LFA model circuit in the VGE *LiveSPICE* environment for analyzing time delays.
The designations of the circuit elements correspond to those adopted in GOST 2.710-811[23]

[23] GOST 2.710-81. Interstate Standard. *Unified system for design documentation. Alpha-numerical designations in electrical diagrams*. Moscow: Izd. Standartov; 1985 (in Russ.).

The time delays resulting from varying the sampling frequency, F_S, within a digital twin of the power amplifier audio module are presented in Table 3, with a constant buffer size of $N = 2048$.

Table 3. Outcomes of examining time delays in a power amplifier's audio module by simulating various sampling frequencies within a digital twin model while maintaining a constant buffer size

VGE environment, data I/O protocol							
LiveSPICE, ASIO		*Simulink*, ASIO		*Simulink*, DirectSound		*Simulink*, WASAPI	
F_S, kHz	τ_{DLY}, ms	F_S, kHz	τ_{DLY}, ms	F_S, kHz	τ_{DLY}, ms	F_S, kHz	τ_{DLY}, ms
44.1	100	44.1	194	44.1	398	44.1	220
48	96	48	182	48	384	48	205
88.2	95	88.2	146	88.2	362	88.2	182
96	95	96	137	96	351	96	157

The time delays resulting from varying buffer size N within the LFA audio module using a digital twin model and maintaining a constant sampling frequency of $F_S = 44.1$ kHz are given in Table 4.

Table 4. The outcomes of examining time delays in a power amplifier's audio module by simulating various buffer sizes while maintaining a constant sampling frequency

VGE environment, data I/O protocol							
LiveSPICE, ASIO		*Simulink*, ASIO		*Simulink*, DirectSound		*Simulink*, WASAPI	
N	τ_{DLY}, ms	N	τ_{DLY}, ms	N	τ_{DLY}, ms	N	τ_{DLY}, ms
8	7	8	100	8	–	8	110
16	7.2	16	100.4	16	–	16	110.2
32	7.4	32	101.2	32	–	32	111.5
64	8	64	102.8	64	–	64	113.3
128	11.6	128	105.2	128	–	128	116.5
256	17.8	256	111.2	256	262	256	118.6
512	29.4	512	124	512	297	512	130
1024	53.2	1024	147	1024	310	1024	150
2048	100	2048	194	2048	398	2048	210

The inclusion of the SPICE circuit results in a doubling of the time delays in comparison to those observed in a direct pass circuit of an operational amplifier without a SPICE model. This increase is attributed to the additional processing time required by the VGE environment software block required to handle the SPICE circuit [10, 14, 15]. Notably, the VGE *LiveSPICE* system exhibits the shortest delay due to the built-in algorithm for calculating the differential equations of the SPICE model just once. Although SPICE integrated algorithm simplifies integral-differential calculations for circuits, its fixed nature hinders its flexibility and restricts the use of pre-built SPICE component libraries from manufacturers as contrasted with the broader capabilities of the Simscape library. Therefore, the formula (2) for hardware delays is presented as follows:

$$T_{INTF} = T_{DAC} + T_{ADC} + T_{Buff.} + T_{Res.} + T_{SR} - T_{VGE}, \quad (3)$$

where T_{VGE} is the time delay introduced by the VGE environment block.

Due to the absence of built-in time measurement tools within the VGE *LiveSPICE* platform, further investigation is conducted in the VGE *Simulink* environment. This research utilizes a custom-developed cyber-physical emulation plugin designed to interact with the Simscape library. Given the ASIO protocol's minimal latency and its widespread use in audio technology [16–18], subsequent research focuses solely on this protocol.

3. MEASURING TIME DELAYS INTRODUCED BY THE SPICE EMULATION BLOCK WITHIN THE *SIMULINK* ENVIRONMENT

For assessing the time delays introduced by the SPICE emulation block in the VGE environment, the built-in *Simulink Profiler*[24] is employed. This tool initiates the simulation and generates a report that may be used to analyze the time delays introduced by the VGE block, which serves as the foundation for the audio module circuit. The window for this tool is shown in Fig. 6.

The time delays introduced by the SPICE emulation block can be evaluated using the obtained parameters as follows:

$$T_{VGE} = \frac{\text{Self Time}}{\text{Number of calls}}, \quad (4)$$

where Self Time is the *Simulink* model operating time, including time for generating the *Simulink Profiler* report,

[24] Technical documentation of *Simulink Profiler*. https://www.mathworks.com/help/simulink/slref/simulinkprofiler.html. Accessed February 20, 2025.

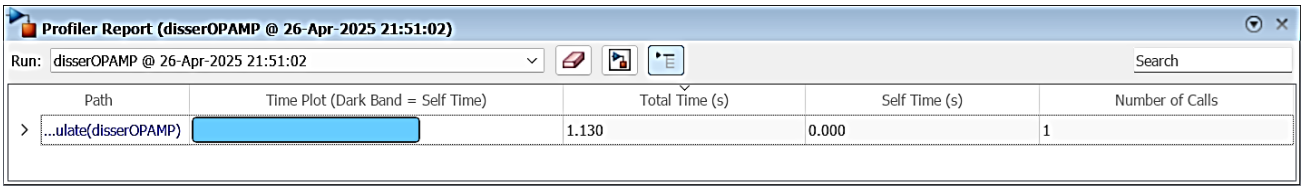


Fig. 6. Window of the built-in *Simulink Profiler* tool with the plugin for cyber-physical emulation of signal audio modules

while Number of Calls is the number of model calls made during Self Time [19, 20].

Tables 5 and 6 present the results of analyzing the time delays introduced by SPICE simulations conducted within *Simulink* for both a direct signal transmission circuit and a circuit utilizing an OA-based LFA.

Table 5. The outcomes of examining the time delays introduced by the SPICE simulation within *Simulink* for the direct signal transmission circuit

Simulink Profiler runtime, s	SPICE simulation time for VGE, ms	SPICE emulation time for VGE with electrical schematic simulation, ms
1	1.2	1.8
2	1.2	1.8
3	1.2	1.8
4	1.2	1.8
5	1.3	1.95

The results presented in Table 5 show that the time delays introduced by the emulation block, taking into account the cyber-physical SPICE emulation of the signal audio module, experience a 1.5-fold augmentation.

Since the VGE *Simulink* environment lacks tools for assessing CPU cycle consumption, a thorough examination of efficiency is impossible. The reported research employed a 64-bit Intel® Core™ i5-10500T (Intel, USA) processor functioning at a frequency of 2.3 GHz.

CONCLUSIONS

The paper presents a novel approach for emulating analog audio systems within a cyber-physical SPICE environment. It proposes an "interface circuit" that utilizes the UMC1820 software/hardware audio interface. The study delves into the time delays inherent in the software/hardware interface of VGE hardware and software systems. Results indicate that the VGE *LiveSPICE* system exhibits the lowest latency (7 ms). This can be attributed to its unique algorithm that calculates SPICE model differential equations only once. Nevertheless, this method lacks flexibility and is incompatible with the utilization of existing SPICE libraries.

Conversely, the VGE *Simulink* system, when enhanced with a plugin for emulating cyber-physical circuits, enables the incorporation of code description of completed PCB models, albeit at the cost of extended time delays (exceeding 100 ms). The *Simulink Profiler* tool is employed to evaluate the computational time delays within the VGE system.

For optimal performance when using the cyber-physical emulation method for audio signal modules, it is advisable to fine-tune the software/hardware interface. This involves hardware calibration using the analog potentiometers integrated into the ADC and software calibration within the VGE system by adjusting time delays through the "Delay" block. Therefore, the optimal configuration for using the cyber-physical emulation method with SPICE models in real-world applications involves a sampling frequency of 44.1 kHz, a buffer size ranging from 512 to 1024 samples, and the ASIO data I/O protocol.

Authors' contributions

N.R. Levchenko developed a plugin for cyber-physical simulation, conducted experiments, and processed findings.

M.S. Kostin suggested a cyber-physical approach to analyzing audio signal modules, offering guidance on conducting and interpreting experimental findings.

REFERENCES

1. Çilingiroğlu Uğur. *Analog Integrated Circuit Design by Simulation: Techniques, Tools, and Methods*. 1st ed. New York: McGraw-Hill Education. 2019. 576 p.
2. Anumba C.J., Akanmu A., Yuan X., Kan C. Cyber–physical systems development for construction applications. *Front. Eng. Manag.* 2021;8(1):72–87. https://doi.org/10.1007/s42524-020-0130-4
3. Levchenko N.R., Kostin M.S. Stand for cyberphysical prototyping of signal audio modules. In: *Actual Problems and Prospects of Development of Radio Engineering and Infocommunication Systems* (*RADIOINFOCOM-2024*): *Proceedings of the 8th International Scientific and Practical Conference*. Moscow: RTU MIREA; 2024. P. 507–511 (in Russ.).
4. Levchenko N.R. Plugin for cyberphysical prototyping of signal audio modules. In: *Fundamental, Exploratory, Applied Research and Innovative Projects: Proceedings of the 3rd National Scientific and Practical Conference*. Moscow. 2024. P. 344–350 (in Russ.). https://elibrary.ru/omggfz
5. Levchenko N.R. Software interfaces for audio data transmission for cyber-physical SPICE-signal prototyping of audio modules. In: *Fundamental, Exploratory, Applied Research and Innovative Projects: Proceedings of the National Scientific and Practical Conference*. Moscow. 2023. P. 394–398 (in Russ.). https://elibrary.ru/pbtjgd
6. Levchenko N.R., Kostin M.S. Cyberphysical prototyping of signal audio modules. In: *Actual Problems and Prospects of Development of Radio Engineering and Infocommunication Systems* (*RADIOINFOCOM-2023*): *Proceedings of the 7th International Scientific and Practical Conference*. Moscow: RTU MIREA; 2023. P. 347–350 (in Russ.). https://www.elibrary.ru/oykued
7. Cordell B. *Designing Audio Power Amplifiers*. Routledge; 2019. 792 p.
8. Petlenko D.B., Yarlykov A.D., Boikov K.A. *Analogo-tsifrovye preobrazovateli signal'nykh audiointerfeisov* (*Analog-to-Digital Converters of Signal Audio Interfaces*). Moscow: Reglet; 2023. 65 p. (in Russ.).
9. Self D. *Small Signal Audio Design*: 4th ed. Focal Press; 2023. 846 p. https://doi.org/10.4324/9781003332985
10. Bennett C.L. *Digital Audio Theory: A Practical Guide*: 1st ed. Focal Press; 2020. 254 p. https://doi.org/10.4324/9780429297144
11. Steiglitz K. A *Digital Signal Processing Primer: with Applications to Digital Audio and Computer Music*. New York, USA: Dover Publications Inc.; 2020. 320 p.
12. Petlenko D.B., Yarlykov A.D., Boikov K.A. *Tsifrovye metody sekvensornoi ekvalizatsii audiosignalov radioakusticheskikh system* (*Digital Methods of Sequencer Equalization of Audio Signals of Radioacoustic Systems*). Moscow: Reglet; 2023. 109 p. (in Russ.).
13. Collins K. *Studying Sound: A Theory and Practice of Sound Design.* London, England: The MIT Press; 2020. 248 p.
14. Kamenov A. *Digital Signal Processing for Audio Applications*: 2nd ed. RecordingBlogs; 2014. 348 p.
15. Cipriani A., Giri M. *Electronic Music and Sound Design: Theory and Practice with Max 8.* V. 2. 3rd ed. ConTempoNet; 2020. 748 p.
16. Kovalgin Yu.A., Vakhitov Sh.Ya. *Akustika* (*Acoustics*). Moscow: Goryachaya liniya – Telekom; 2022. 660 p. (in Russ.).
17. Pirkle W.C. *Designing Audio Effect Plugins in C++: for AAX, AU, and VST3 with DSP Theory*: 2nd ed. New York, USA: Routledge; 2019. 704 p.
18. Gevorsky A.V., Kostin M.S., Boikov K.A. Software-architectural configuration of the multifunctional audio digital signal processor module for signal mediatesting of audio devices. *Russ. Technol. J.* 2024;12(1):30–58. https://doi.org/10.32362/2500-316X-2024-12-1-30-58
19. Oney W. *Programming the Microsoft® Windows® Driver Model*: 2nd ed. Microsoft Press; 2002. 880 p.
20. Yosifovich P. *Windows Kernel Programming*. Independently published; 2021. 461 p.

СПИСОК ЛИТЕРАТУРЫ

1. Çilingiroğlu Uğur. *Analog Integrated Circuit Design by Simulation: Techniques, Tools, and Methods*. 1st ed. New York: McGraw-Hill; 2019. 576 p.
2. Anumba C.J., Akanmu A., Yuan X., Kan C. Cyber—physical systems development for construction applications. *Front. Eng. Manag.* 2021;8(1):72–87. https://doi.org/10.1007/s42524-020-0130-4
3. Левченко Н.Р., Костин М.С. Стенд киберфизического прототипирования сигнальных аудиомодулей. В сб.: *Актуальные проблемы и перспективы развития радиотехнических и инфокоммуникационных систем* («*Радиоинфоком-2024*»): *Сборник научных статей по материалам VIII Международной научно-практической конференции*. М.: РТУ МИРЭА; 2024. С. 507–511.
4. Левченко Н.Р. Плагин для киберфизического прототипирования сигнальных аудиомодулей. В сб.: *Фундаментальные, поисковые, прикладные исследования и инновационные проекты*: *Сборник трудов III Национальной научно-практической конференции*. Москва. 2024. С. 344–350. https://elibrary.ru/omggfz
5. Левченко Н.Р. Программные интерфейсы передачи аудиоданных для киберфизического SPICE-сигнального прототипирования аудиомодулей. В сб.: *Фундаментальные, поисковые, прикладные исследования и инновационные проекты*: *Сборник трудов Национальной научно-практической конференции*. Москва. 2023. С. 394–398. https://elibrary.ru/pbtjgd

6. Левченко Н.Р., Костин М.С. Киберфизическое прототипирование сигнальных аудиомодулей. В сб.: *Актуальные проблемы и перспективы развития радиотехнических и инфокоммуникационных систем* («*Радиоинфоком-2023*»): *Сборник научных статей по материалам VII Международной научно-практической конференции*. М.: РТУ МИРЭА; 2023. С. 347–350. https://www.elibrary.ru/oykued
7. Cordell B. *Designing Audio Power Amplifiers*. Routledge; 2019. 792 p.
8. Петленко Д.Б., Ярлыков А.Д., Бойков К.А. *Аналого-цифровые преобразователи сигнальных аудиоинтерфейсов.* М.: Реглет; 2023. 65 с.
9. Self D. *Small Signal Audio Design:* 4th ed. Focal Press; 2023. 846 p. https://doi.org/10.4324/9781003332985
10. Bennett C.L. *Digital Audio Theory: A Practical Guide:* 1st ed. Focal Press; 2020. 254 p. https://doi.org/10.4324/9780429297144
11. Steiglitz K. *A Digital Signal Processing Primer: with Applications to Digital Audio and Computer Music.* New York, USA: Dover Publications Inc.; 2020. 320 p.
12. Петленко Д.Б., Ярлыков А.Д., Бойков К.А. *Цифровые методы секвенсорной эквализации аудиосигналов радиоакустических систем.* М.: Реглет; 2023. 109 с.
13. Collins K. *Studying Sound: A Theory and Practice of Sound Design.* London, England: The MIT Press; 2020. 248 p.
14. Kamenov A. *Digital Signal Processing for Audio Applications*: 2nd ed. RecordingBlogs; 2014. 348 p.
15. Cipriani A., Giri M. *Electronic Music and Sound Design: Theory and Practice with Max 8.* V. 2. 3rd ed. ConTempoNet; 2020. 748 p.
16. Ковалгин Ю.А., Вахитов Ш.Я. *Акустика.* М.: Горячая линия – Телеком; 2022. 660 с.
17. Pirkle W.C. *Designing Audio Effect Plugins in C++: for AAX, AU, and VST3 with DSP Theory*: 2nd ed. New York, USA: Routledge; 2019. 704 p.
18. Геворский А.В., Костин М.С., Бойков К.А. Программно-архитектурная конфигурация многофункционального ADSP-модуля сигнального медиатестирования аудиоустройств. *Russ. Technol. J.* 2024;12(1):30–58. https://doi.org/10.32362/2500-316X-2024-12-1-30-58
19. Oney W. *Programming the Microsoft® Windows® Driver Model*: 2nd ed. Microsoft Press; 2002. 880 p.
20. Yosifovich P. *Windows Kernel Programming*. Independently published; 2021. 461 p.

About the Authors

Nikita R. Levchenko, Assistant, Department of Radio Wave Processes and Technologies, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: Levchenko_n@mirea.ru. RSCI SPIN-code 5126-8830, https://orcid.org/0009-0002-1749-1451

Mihail S. Kostin, Dr. Sci. (Eng.), Associate Professor, Head of the Department of Radio Wave Processes and Technologies, Deputy Director, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: kostin_m@mirea.ru. Scopus Author ID 57208434671, RSCI SPIN-code 5819-2178, http://orcid.org/0000-0002-5232-5478

Об авторах

Левченко Никита Романович, ассистент, кафедра радиоволновых процессов и технологий, Институт радиоэлектроники и информатики, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: Levchenko_n@mirea.ru. SPIN-код РИНЦ 5126-8830, https://orcid.org/0009-0002-1749-1451

Костин Михаил Сергеевич, д.т.н., доцент, заведующий кафедрой радиоволновых процессов и технологий, заместитель директора Института радиоэлектроники и информатики, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: kostin_m@mirea.ru. Scopus Author ID 57208434671, SPIN-код РИНЦ 5819-2178, http://orcid.org/0000-0002-5232-5478

Translated from Russian into English by K. Nazarov
Edited for English language and spelling by Thomas A. Beavitt

Modern radio engineering and telecommunication systems

Современные радиотехнические и телекоммуникационные системы

UDC 621.391.825
https://doi.org/10.32362/2500-316X-2025-13-5-63-74
EDN QHZTAK

RESEARCH ARTICLE

Non-fluctuation interference rejection using an adaptive filter based on spectrum envelope analysis

Georgy V. Konyashkin [@], Gennady V. Kulikov

MIREA – Russian Technological University, Moscow, 119454 Russia
[@] Corresponding author, e-mail: konyashkin@mirea.ru

• **Submitted:** 15.04.2025 • **Revised:** 29.05.2025 • **Accepted:** 23.07.2025

Abstract

Objectives. The rapid advancement of wireless technologies, including IoT and 5G/6G, is accompanied by an increase in the overall level of electromagnetic interference. This sets engineers the task of developing effective methods to suppress such interference, including especially challenging non-fluctuating interference of various kinds. The study aims to implement and analyze the effectiveness of a non-fluctuation interference rejection method using an adaptive filter based on spectrum envelope analysis.

Methods. Mathematical modeling, spectral analysis, and adaptive filtering methods are used in the work. The described approach is based on spectrum envelope extraction for identification and subsequent suppression of non-fluctuation interference.

Results. The effectiveness of an adaptive algorithm for suppressing non-fluctuation interference based on the analysis of the spectrum envelope has been demonstrate. This algorithm can be used as a means for isolating the envelope of the interference spectrum to enable the formation of the amplitude-frequency response of the notch filter in real time. Processing methods for three types of non-fluctuation interference were implemented and tested: harmonic, frequency-shift keying (FSK), and phase-shift keying (PSK). A signal with quadrature amplitude modulation forms a useful signal for the purposes of the study. The experimental results demonstrate the good efficiency of the proposed method. The developed adaptive notch filter based on spectrum envelope analysis is highly effective in combating harmonic interference to achieve energy gains of 8–9 dB depending on the relative intensity of interference. Notably, even as interference intensifies, the filter effectiveness persists, albeit with a slight reduction. The algorithm functions effectively under exposure to narrowband FSK and PSK interference.

Conclusions. The proposed adaptive algorithm for suppressing fluctuation interference based on spectrum envelope analysis is optimally effective in the presence of harmonic interference within the communication channel, but less effective in the presence of more broadband interference. The study is of practical importance for digital communication systems, where high noise immunity is required in a complex electromagnetic environment.

Keywords: quadrature amplitude modulation, non-fluctuation interference, spectrum envelope, adaptive filtering, rejection, bit error rate, noise immunity

For citation: Konyashkin G.V., Kulikov G.V. Non-fluctuation interference rejection using an adaptive filter based on spectrum envelope analysis. *Russian Technological Journal.* 2025;13(5):63–74. https://doi.org/10.32362/2500-316X-2025-13-5-63-74, https://www.elibrary.ru/QHZTAK

Financial disclosure: The authors have no financial or proprietary interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Режекция нефлуктационных помех с помощью адаптивного фильтра на основе анализа огибающей спектра

Г.В. Коняшкин [@], Г.В. Куликов

МИРЭА – Российский технологический университет, Москва, 119454 Россия
[@] Автор для переписки, e-mail: konyashkin@mirea.ru

• **Поступила:** 15.04.2025 • **Доработана:** 29.05.2025 • **Принята к опубликованию:** 23.07.2025

Резюме

Цели. Активное развитие беспроводных технологий, включая IoT (Internet of Things, интернет вещей) и 5G/6G, сопровождается ростом уровня электромагнитных помех, что ставит перед инженерами задачу разработки эффективных методов их подавления. Особую сложность представляют нефлуктуационные помехи разного рода. Цель данного исследования заключается в реализации и анализе эффективности метода режекции нефлуктуационных помех с использованием адаптивного фильтра, основанного на анализе огибающей спектра.

Методы. В работе использованы методы математического моделирования, спектрального анализа и адаптивной фильтрации. Предложен подход, основанный на выделении огибающей спектра для идентификации и последующего подавления нефлуктуационных помех.

Результаты. Проведено исследование эффективности адаптивного алгоритма подавления нефлуктуационных помех на основе анализа огибающей спектра, который позволяет выделять огибающую спектра помехи, что обеспечивает формирование амплитудно-частотной характеристики режекторного фильтра в реальном времени. В ходе исследования реализованы и протестированы методы обработки для 3 типов нефлуктуационных помех: гармонической, частотно-манипулированной (ЧМ) и фазоманипулированной (ФМ). В качестве полезного сигнала использован сигнал с квадратурной амплитудной модуляцией. Экспериментальные результаты демонстрируют хорошую эффективность предложенного метода, Разработанный адаптивный режекторный фильтр на основе анализа огибающей спектра обладает высокой эффективностью при борьбе с гармонической помехой: энергетический выигрыш в зависимости от относительной интенсивности помехи может составлять до 8–9 дБ. При увеличении количества помех эффективность фильтра сохраняется, хотя и несколько снижается. Алгоритм при определенных условиях работоспособен в условиях воздействия узкополосных ЧМ- и ФМ-помех.

Выводы. Предложенный адаптивный алгоритм подавления нефлуктуационных помех на основе анализа огибающей спектра наиболее эффективен при наличии в канале связи гармонических помех и менее эффективен при наличии более широкополосных помех. Работа имеет практическую значимость для систем цифровой связи, где требуется высокая помехоустойчивость в условиях сложной электромагнитной обстановки.

Ключевые слова: квадратурная амплитудная модуляция, нефлуктуационные помехи, огибающая спектра, адаптивная фильтрация, режекция, вероятность битовой ошибки, помехоустойчивость

Для цитирования: Коняшкин Г.В., Куликов Г.В. Режекция нефлуктационных помех с помощью адаптивного фильтра на основе анализа огибающей спектра. *Russian Technological Journal*. 2025;13(5):63–74. https://doi.org/10.32362/2500-316X-2025-13-5-63-74, https://www.elibrary.ru/QHZTAK

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

The rapid advancement of wireless technologies, such as the Internet of Things (IoT) and 5G/6G, is accompanied by an increase in the overall level of electromagnetic interference. This poses a challenge to engineers in terms of developing effective methods for their suppression. The suppression of non-fluctuating interference turns out to be particularly challenging. Such interference includes harmonic signals, as well as frequency and phase modulation signals, which may be caused by either accidental technical violations of radio regulations or intentional actions. Such interference significantly reduces signal-to-interference-plus-noise ratio (SINR), hindering information decoding and elevating the probability of errors within communication systems [1–7].

Conventional suppression methods, including adaptive filters utilizing LMS[1]/NLMS[2] algorithms or fast Fourier transform (FFT) approaches, are often ineffective [8–12]. This insufficiency stems from their restricted capacity to adapt to transient interference and high processing computational costs. Furthermore, methods relying on initial spectrum analysis necessitate precise knowledge of the interference frequency response, which represents a particular challenge when faced with dynamically changing interference.

The present work introduces a novel method for leveraging a priori information on the spectral envelope of the useful signal. After extracting a spectral interference mask, a notch filter for eliminating the identified frequency components is synthesized [13]. A significant benefit of this algorithm is its adaptability, which allows it to be operated in the absence of a priori information on interference characteristics and in real-time. To evaluate its versatility, the algorithm is subjected to tests involving three distinct categories of non-fluctuation interference. The experimental results demonstrate that the developed adaptive filter effectively minimizes interference levels while maintaining a minimal effect on the useful signal.

This study is relevant due to the growing requirements for noise immunity of modern radio systems, especially those arising in the context of the rise of IoT technologies and autonomous devices, where reliable communication is crucial for both safety and proper function. The findings of this study can be applied to improve the performance of telecommunications equipment, military communication networks, and industrial automation systems, enabling them to operate effectively in challenging electromagnetic environments.

INTERFERENCE REJECTION METHOD

The adaptive filter implemented during the study, which is designed based on spectrum envelope analysis, functions as follows:

1. Spectral processing. The received signal-interference mixture is subjected to FFT spectral analysis. The conversion of the signal from the time domain to the frequency domain provides a basis for obtaining a spectral picture for further analysis.
2. Spectrum envelope extraction. The spectrum envelope of the received mixture is calculated by extrapolating values of local maximums of the spectrum over a given range of points. This enables the clear identification and isolation of individual frequency components within a mixture of signals and interference.
3. Spectrum envelope analysis. To identify the frequency components that require filtering, the spectrum envelope of the received signal is examined. For this, the a priori known (since the receiving side knows the main parameters of the signal – in particular, the modulation method and the information transmission rate) spectrum reference envelope of the useful signal is subtracted from the spectrum envelope of the received mixture. This allows extraction of the interference spectrum envelope, whose parameters can then be utilized in configuring the filtering system.

[1] The Least Mean Squares (LMS) algorithm is an adaptive filtering technique used to adjust the filter coefficients so as to minimize the mean square error between the filter output and the desired signal.

[2] The Normalized Least Mean Squares (NLMS) algorithm is an adaptive filtering method used to optimize filter parameters in real time. It extends the basic LMS algorithm by normalizing the adaptation step, which increases its stability and efficiency when working with time-varying signals.

4. Synthesis of the notch filter. The derived interference spectrum envelope can be used for the real-time generation of an amplitude-frequency response for a customizable notch filter, which filters the mixture of signal and interference. This adaptable notch filter can be designed as a multi-band system with customizable gain levels across different frequency bands [13], or as a filter utilizing a synthesized amplitude-frequency response, the latter approach being implemented in this study.

A functional diagram of the adaptive filter implementing this algorithm is shown in Fig. 1.

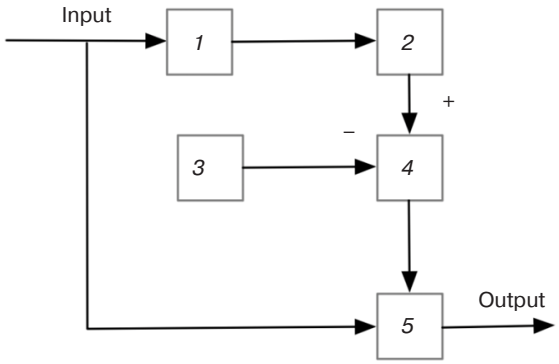


Fig. 1. Functional diagram of the adaptive filter based on the spectrum envelope analysis: *1* is a Fourier converter; *2* is a spectrum envelope detector; *3* is a reference spectrum envelope of the useful signal; *4* is an adder; and *5* is a notch filter with synthesized amplitude-frequency response

ADAPTIVE FILTERING COMPUTER SIMULATION

An adaptive filter based on spectrum envelope analysis is simulated in *MATLAB/Simulink* (trial version) and *Scilab/Xcos* for a 16-Quadrature Amplitude Modulation (16-QAM) signal (Fig. 2). The results of the study can be easily adapted to other signals with stationary spectrum.

Within the "M QAM Modulation" section, a useful signal of a given dimension is generated, after which an additive white Gaussian noise is superimposed in the channel and added to the selected type of non-fluctuation interference generated in the "Generation of various types of non-fluctuation interference" section. Then, if necessary, transfer to the carrier frequency is performed, after which the resulting mixture of signal and interference is branched into 2 sections of demodulation and BER calculation for processing with and without an adaptive filter. This also falls into the subsystem for isolating the envelope of the interference spectrum detailed in Fig. 3.

The spectrograms obtained as a result of the circuit operation clearly reveal the principle of operation and validate the functionality of the algorithm illustrated in Fig. 4.

The three distinct interference generators created for research purposes encompass harmonic interference (ranging from 1 to 4 harmonics), frequency-shift keying (FSK), and phase-shift keying (PSK) types. The diagram's upper section illustrates an algorithm designed to isolate the interference spectrum envelope in adherence to the principles outlined in Fig. 1. Data from the analyzed envelope of the interference spectrum is processed with the Yule–Walker function[3] to create IIR filters that possess a predetermined frequency response. This design process, which relies on the Yule–Walker equations [14–16], results in a filter with a specific transfer function:

$$H(z) = \frac{B(z)}{A(z)} = \frac{b_0 + b_1 z^{-1} + ... + b_{n_b} z^{-n_b}}{1 + a_1 z^{-1} + ... + a_{n_a} z^{-n_a}},$$

where n_a and n_b are the degrees of the denominator and numerator, respectively, while z is a complex variable of the Z-transform $(z = e^{j\omega_n})$ that bridges the filter's difference equation and its transfer function; $\omega = \frac{2\pi n}{N}$.

Through an iterative process, the interference spectrum envelope algorithm aims to minimize the mean square error between the desired and actual frequency response.

The main stages of the algorithm are the following:

1. Defining the desired frequency response. The user defines the vectors of frequency f and corresponding amplitudes m, specifying the desired characteristic. Frequencies are normalized relative to the Nyquist frequency ($0 \le f \le 1$).
2. Calculating the autocorrelation function. The autocorrelation function $r(k)$ describes how the signal correlates with itself at various time delays k. The desired frequency response is transformed into an autocorrelation function using the inverse Fourier transform as expressed by:

$$r(k) = \frac{1}{N} \sum_{n=0}^{N-1} |H_d(e^{j\omega_n})|^2 \, e^{j2\pi kn/N},$$

where $|H_d(e^{j\omega_n})|^2$ is the squared amplitude of the desired characteristic, N is the number of sampling points, and k is the delay index.

[3] *Signal Processing Toolbox User's Guide*. Copyright 1988–2002 by The MathWorks, Inc.

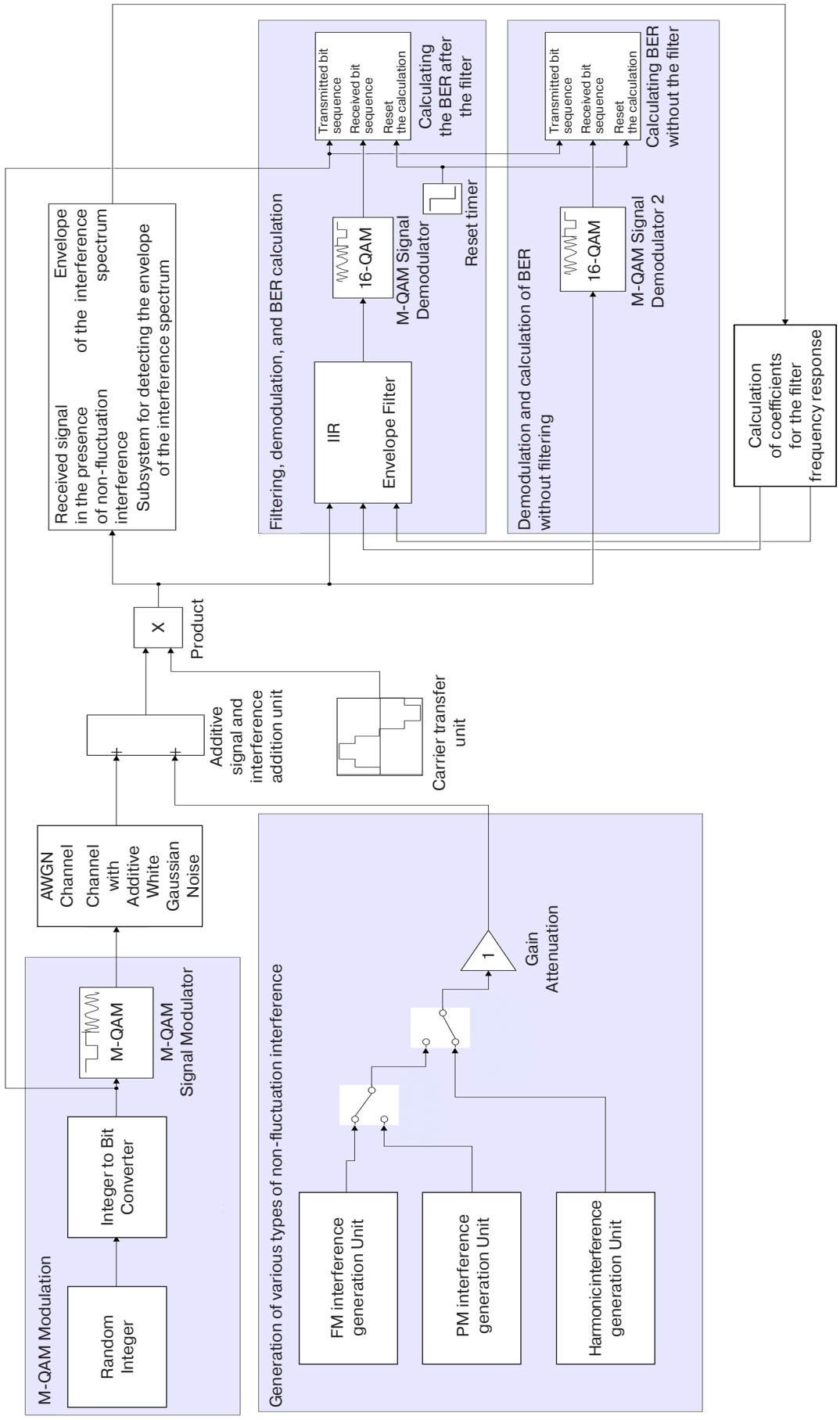


Fig. 2. Functional diagram for investigating the influence of different types of non-fluctuation interference. IIR is infinite impulse response and BER is bit error rate

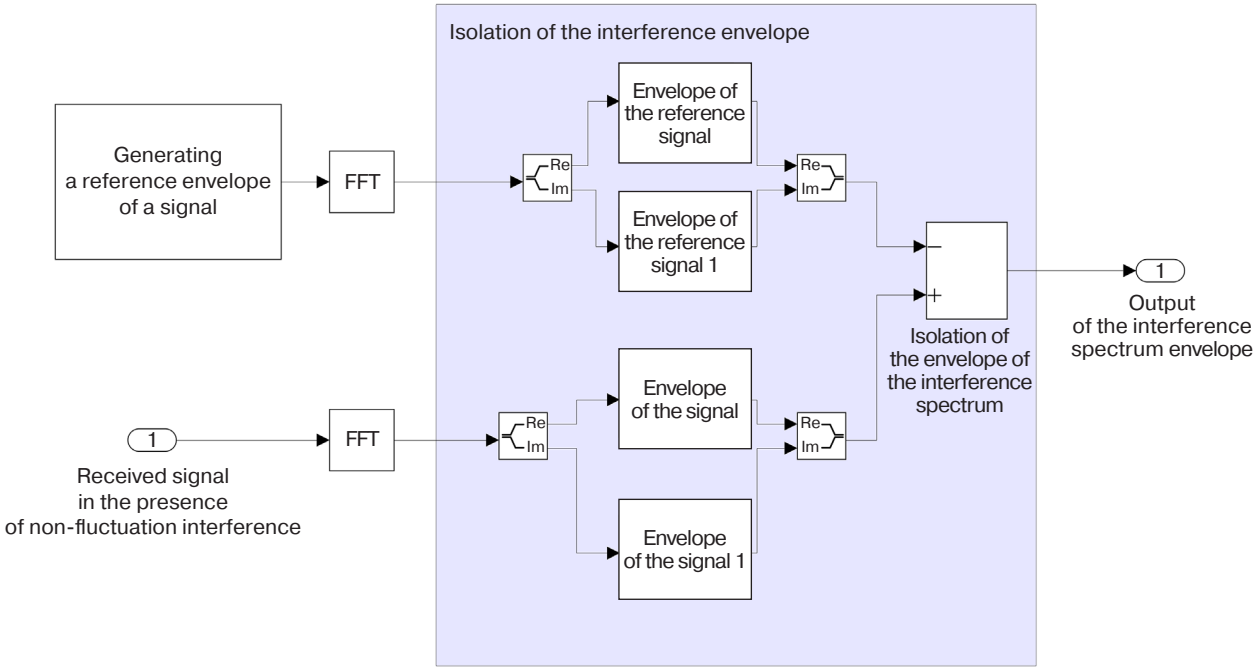


Fig. 3. Subsystem for isolating the envelope of the interference spectrum

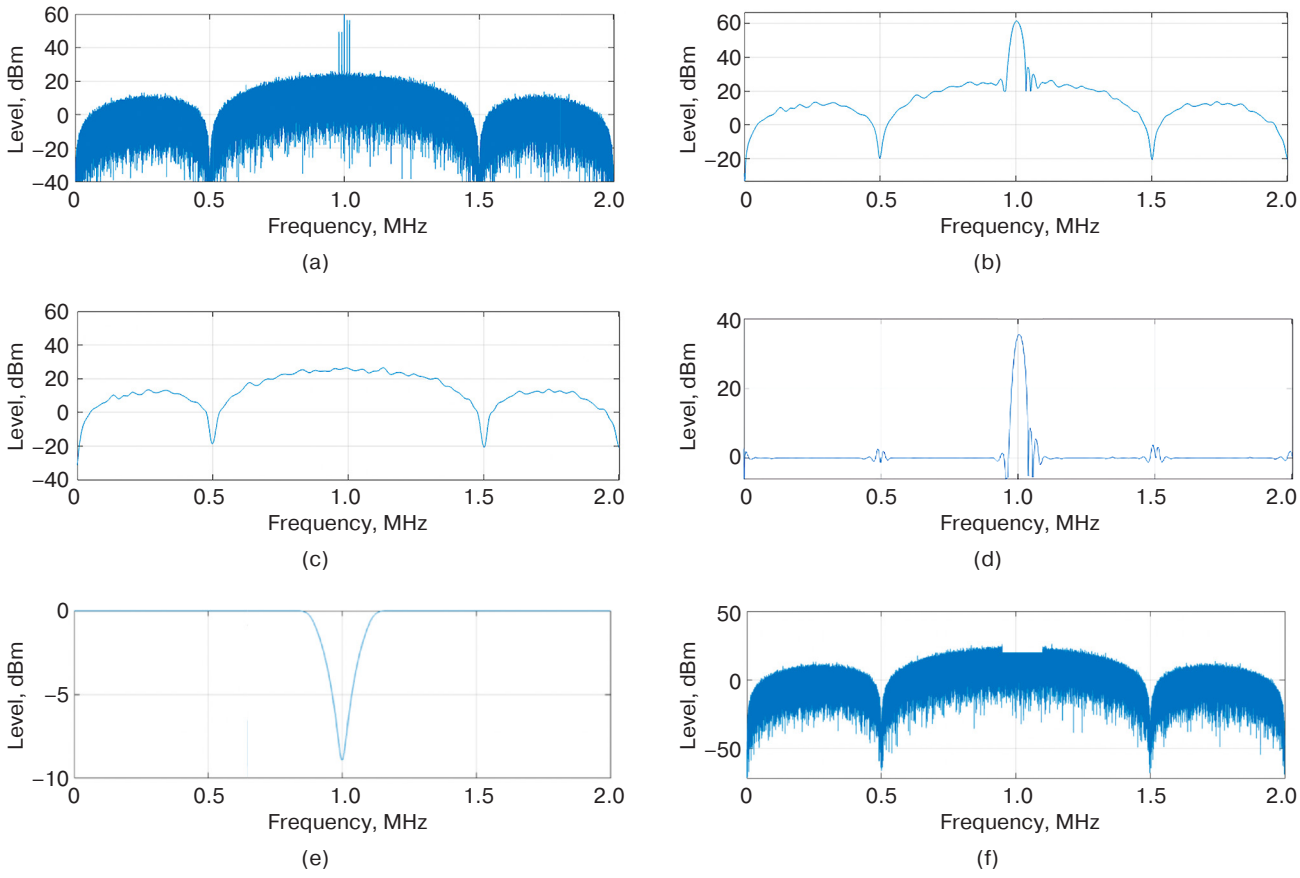


Fig. 4. Spectrograms illustrating the principle of the algorithm:
(a) the spectrum of the received signal with interference; (b) envelope of spectrum of received signal with interference; (c) signal envelope reference; (d) difference of envelopes of received signal with interference and reference envelope of signal spectrum; (e) synthesized frequency response of the notch filter; and (f) a spectrum of the filtered signal. The level in this case is the relative unit of signal power, expressed in decibels with respect to 1 mW

3. Solving the Yule–Walker equations for the denominator $A(z)$ of the transfer function. The $n \times n$ autocorrelation matrix **R** is formed, depending on the required filter order. Given a filter order of n, the denominator coefficients $\mathbf{a} = [a_1, a_2, \ldots, a_n]^T$ are derived from the following system:

$$\mathbf{R} \cdot \mathbf{a} = -r,$$

where $R_{ij} = r(|i - j|)$ are elements of the autocorrelation matrix; $r = [r(1), r(2), \ldots, r(n)]^T$. Next, the denominator is calculated in the form of $A(z) = 1 + a_1 z^{-1} + \ldots a_n z^{-n}$.

4. Calculating the coefficients of the numerator $B(z)$ of the transfer function. After committing $A(z)$, the coefficients of the numerator $B(z)$ are determined by the least-squares method, minimizing the error, as follows:

$$\sum_{i=0}^{L-1} \left| H_d(e^{j\omega_n}) - \frac{B(e^{j\omega_i})}{A(e^{j\omega_i})} \right|^2,$$

where $A(z)$ is the denominator already found and L is the number of frequency points.

Then, for simplicity, a transition to a linear form is performed when calculating the numerator $B(z)$, as follows:

$$B(z) \approx H_d(e^{j\omega_i}) A(z).$$

5. Iterative refinement. At this point, the algorithm may re-adjust $A(z)$ and $B(z)$ to enhance the alignment with the desired frequency response.

SIMULATION RESULTS

Harmonic interference

For harmonic interference, rejection efficiency is studied taking into account the amount of interference simultaneously present in the communication channel, the relative amplitude of μ interference (interference amplitude divided by the average amplitude of the useful signal), and the signal-to-noise ratio (SNR). The resulting BER is then employed to gauge the rejection efficiency.

The results of the study into the impact of harmonic interference levels on filter efficiency are presented in Table.

The obtained values plotted in Fig. 5 illustrate the filtering efficiency depending on the amount of harmonic interference.

The graph demonstrates that the adaptive filtering method, which is most efficient when dealing with a small amount of harmonic interference, maintains its effectiveness even as the interference increases.

The BER vs. SNR relationships at different relative intensities of one harmonic interference μ for two cases, without filtering and using the developed adaptive notch filter, are shown in Fig. 6.

Table. Study results depending on the amount of harmonic interference at SNR = 13 dB, μ = 1

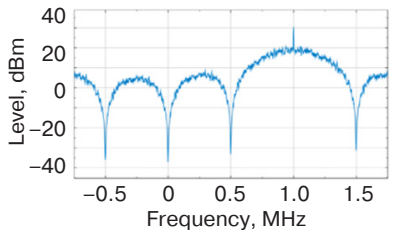
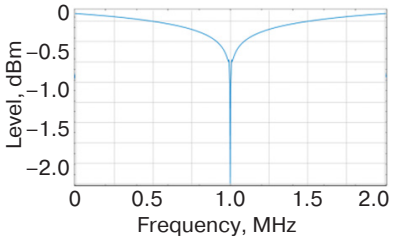
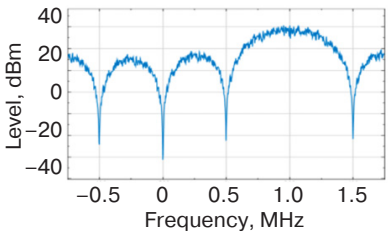
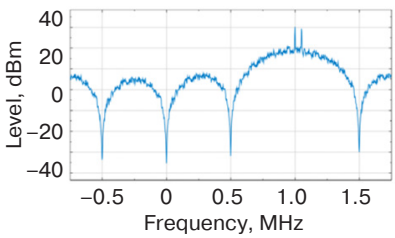
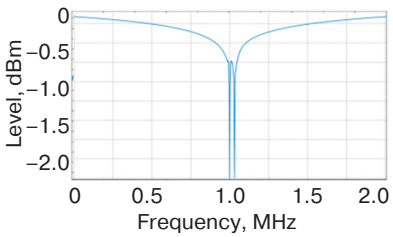
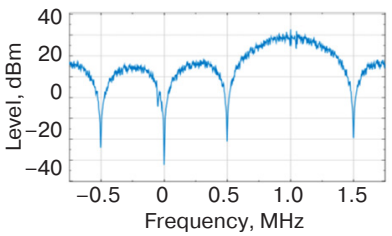
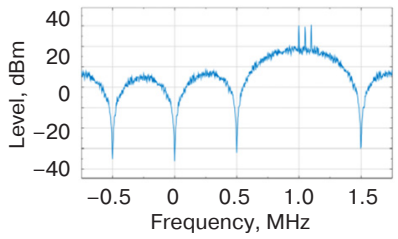
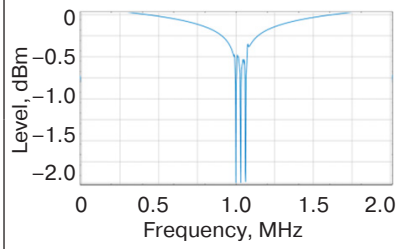
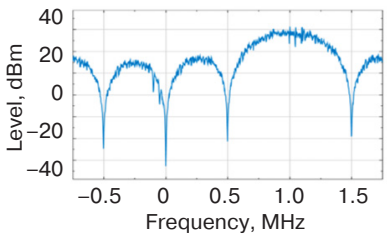
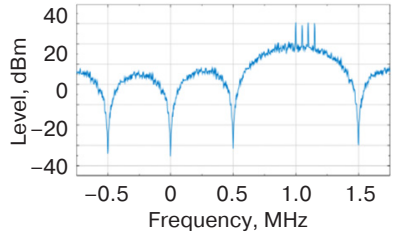
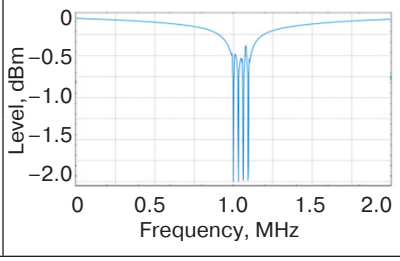
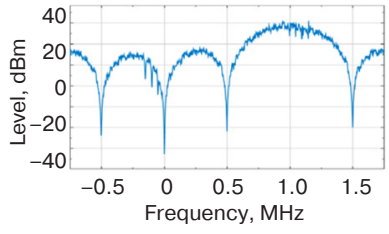
Spectrum before filtration	Filter frequency response	Spectrum after filtration	BER with/without filtration
			0.01059 0.09377
			0.01294 0.04491

Table. Continued

Spectrum before filtration	Filter frequency response	Spectrum after filtration	BER with/without filtration
			0.06486
			0.14781
			0.08102
			0.13731

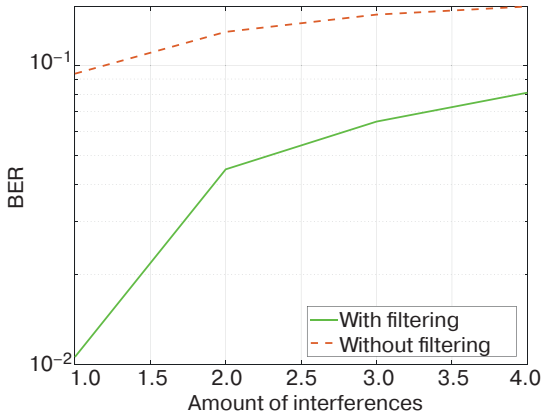


Fig. 5. BER vs. the amount of interference

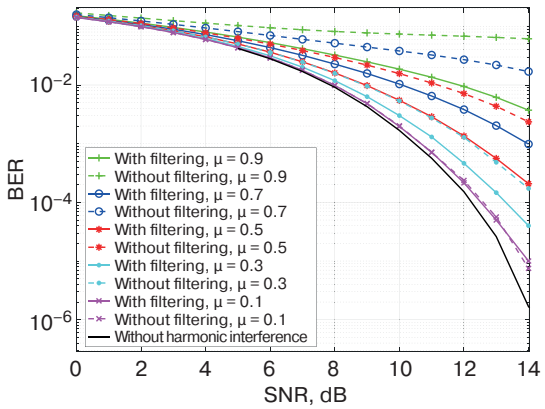


Fig. 6. BER vs. SNR using a notch filter based on the spectrum envelope analysis and without a filter when receiving 16-QAM signals against harmonic interference with different relative intensity μ

It can be concluded from the graphs that the use of the developed filter is efficient at any relative amplitude of interference. For μ = 0.9, the energy gain is on the order of 8–9 dB at BER = 10^{-2} and for μ = 0.5, it is 3 dB at BER = 10^{-3}.

FSK interference

The width of the FSK interference spectrum depends on its transmission rate (keying frequency) V_{inter} and frequency deviation F_{dev}. Therefore, the BER dependence on different deviation $F_{dev\ rel}$ normalized relative to the channel signal rate V_s is studied. The interference-to-signal transmission rate ratio is selected from the condition $V_{inter} << V_s$. In this case, the interference effect is concentrated on the main lobe of the QAM signal.

The BER vs. SNR relationships at different deviation $F_{dev\ rel}$ of the FSK interference for two cases, without filtering and using the developed adaptive notch filter, are shown in Fig. 7.

The graph indicates that this adaptive filter is more successful at mitigating harmonic interference as compared to FSK interference. However, it can also be used to partially reduce interference with a low frequency deviation $F_{dev\ rel} \leq 0.01$. For higher deviations, the filter proves ineffective.

This study investigates the impact of FSK interference with varying relative intensity μ at $F_{dev\ rel} = 0.01$, both in the absence of filtration and with the implementation of a synthesized notch filter. The findings are illustrated in Fig. 8.

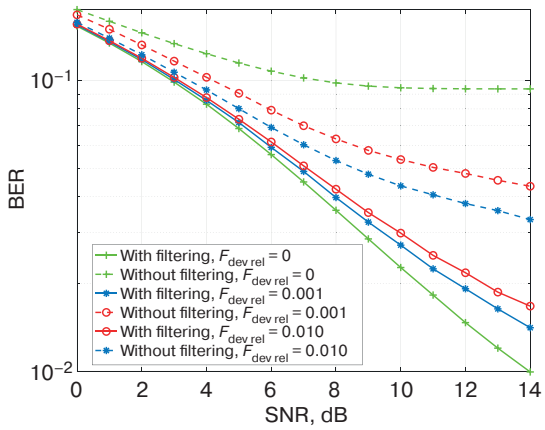


Fig. 7. BER vs. SNR using a notch filter based on the spectrum envelope analysis and without a filter when receiving 16-QAM signals against a background of FSK interference with different relative deviation at μ = 1

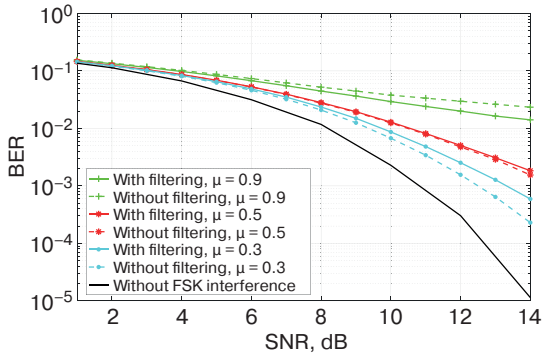


Fig. 8. BER vs. SNR using a notch filter based on spectrum envelope analysis and without a filter when receiving 16-QAM signals against FSK interference with varying relative intensity μ

It can be concluded from the graphs that the use of the developed filter is efficient at a relative interference intensity of $\mu > 0.9$ only. In other cases, the useful signal is rejected. However, under favorable conditions, the energy gain is on the order of 2–4 dB at BER = 10^{-2}.

PSK interference

Since the width of the PSK interference spectrum depends on its transmission rate (keying frequency) V_{inter}, it is the BER dependence on different transmission rate $V_{inter\,rel}$ normalized to the channel signal rate V_s that is investigated.

The dependence of BER on SNR when receiving a 16-QAM signal against PSK interference with a relative amplitude of $\mu = 1$ and a different relative transmission rate $V_{inter\,rel}$ for the case without filtering and that using the developed adaptive notch filter is shown in Fig. 9.

An examination of the generated graphs reveals that the implemented filter is more successful in mitigating harmonic interference as compared to PSK interference. Nevertheless, it also demonstrates the ability to partially suppress interference having a low relative transmission rate $V_{inter\,rel} \le 0.01$. At higher relative transmission rates, the filter is ineffective.

The simulation results of the investigation into the effect of the PSK interference with varying relative intensity μ at $V_{inter\,rel} = 0.01$ are shown in Fig. 10.

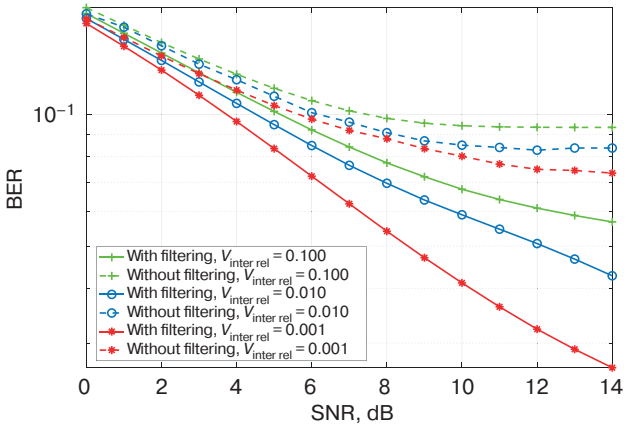


Fig. 9. BER vs. SNR using a notch filter based on spectrum envelope analysis and without filter when receiving 16-QAM signals against PSK interference at different relative transmission rates

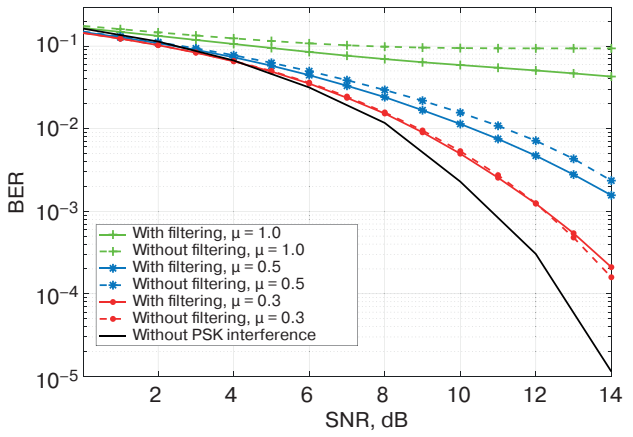


Fig. 10. BER vs. SNR using a notch filter based on the spectrum envelope and without filtering when receiving 16-QAM signals against PSK interference at $V_{inter\,rel}$ = 0.01 and different relative intensity μ

It can be seen that the developed filter, as in the case of FSK-interference, is efficient at high $\mu = 1$, thus providing an energy gain of several decibels. In addition, it provides a small energy gain at average values of $\mu = 0.5$. However, the situation is exacerbated at low $\mu < 0.3$. This is due to the fact that the components of the useful signal are also rejected, which is more noticeable in the latter case. However, for $\mu > 0.5$, there is an energy gain on the order of 2–4 dB and BER = 10^{-2}.

CONCLUSIONS

Based on the study results, the following can be concluded:

1. The developed adaptive notch filter based on spectrum envelope analysis is highly efficient in combating harmonic interference. It can achieve an energy gain of 8–9 dB depending on the interference intensity. With an increase in the amount of interference, the filter efficiency is maintained although at a slightly reduced rate.
2. Initially designed to address harmonic interference, the rejection algorithm demonstrates effectiveness in managing narrowband FSK and PSK interference in specific scenarios.
3. For the FSK interference rejection, the effectiveness of the filter is highly influenced by both the deviation of the interference frequency and its intensity. The filter performs best when dealing with narrowband FSK interference; however, its effectiveness diminishes as the frequency deviation increases. At low frequency deviation ($F_{\text{dev rel}} = 0.01$) and high interference intensity ($\mu > 0.5$), employing the filter can yield an energy gain of 2–4 dB.
4. For PSK interference rejection, the effectiveness of the notch filter is heavily influenced by both the interference rate and its intensity, showing a decline in efficiency as the interference rate rises. However, employing the filter can still yield an energy gain of 2–4 dB at $\mu > 0.5$.
5. The effectiveness of the developed notch filter improves as the relative intensity of FSK and PSK interference increases. However, when the relative intensity is below 0.5, filtering is ineffective due to significant rejection of the useful signal.
6. The effectiveness of the filter hinges on the non-fluctuation interference maintaining a consistent state, which is directly influenced by the iterative synthesis algorithm speed for achieving the desired frequency response.

Authors' contribution
All authors equally contributed to the research work.

REFERENCES

1. Savvateev Yu.I., Nazarov O.V. (Eds.). *Pomekhozashchishchennost' priema diskretnykh signalov* (*Noise Immunity of Reception of Discrete Signals*). Moscow: Radiotekhnika; 2015. 584 p. (in Russ.). ISBN 978-5-93108-094-9
2. Borisov V.I., Zinchuk V.M. *Pomekhozashchishchennost' sistem radiosvyazi. Veroyatnostno-vremennoi podkhod* (*Noise Immunity of Radio Communication Systems. Probabilistic-Temporal Approach*). Moscow: RadioSoft; 2008. 260 p. ISBN 5-93274-011-6 (in Russ.). https://www.elibrary.ru/catzhm
3. Parshutkin A.V., Maslakov P.A. Noise stability of sattelite communication channels with amplitude-phase modulation to exposure to urged unsteady interference. *Voprosy oboronnoi tekhniki. Seriya 16. Tekhnicheskie sredstva protivodeistviya terrorizmu = Military Enginery. Counter-Terrorism Technical Divices. Issue 16.* 2019;11–12:96–101 (in Russ.).
4. Lozhkin K.Yu., Petrov A.V., Mironov V.A., Mikhalev V.V., Prozhetorko S.S. Analytical dependences of bit distortion average probability M-QAM of a signal against harmonic or PSK jamming subject to fading. *Radiotekhnika = Radioengineering.* 2020;84(4–8):27–35 (in Russ.). https://doi.org/10.18127/j00338486-202004(8)-03
5. Kulikov G.V., Nesterov A.V., Lelyukh A.A. Interference immunity of reception of signals with quadrature amplitude shift keying in the presence of harmonic interference. *Zhurnal Radioelektroniki = Journal of Radio Electronics*. 2018;11:2 (in Russ.). https://doi.org/10.30898/1684-1719.2018.11.9
6. Kulikov G.V., Lelyukh A.A., Batalov E.V., Kuzelenkov P.I. Noise immunity of reception of signals with quadrature amplitude modulation in the presence of interference phase-shift keying. *Zhurnal Radioelektroniki = J. Radio Electronics* 2019;7 (in Russ.). Available from URL: http://jre.cplire.ru/jre/jul19/10/text.pdf, https://doi.org/10.30898/1684-1719.2019.7.10
7. Kulikov G.V., Shamshura A.O., Pechenin E.A., Shatalov E.V. Analysis of the noise immunity of receiving signals with quadrature amplitude modulation against the background of frequency-shift keyed interference. *Vestnik Voronezhskogo instituta FSIN Rossii = Vestnik of Voronezh Institute of the Russian Federal Penitentiary Service* 2022;2:9–15 (in Russ.).
8. Widrow B., Stearns S.D. *Adaptive Signal Processing*. Prentice-Hall; 1985. 474 p.
[Widrow B., Stearns S.D. *Adaptivnaya obrabotka signalov* (*Adaptive Signal Processing*): transl. from Engl. Moscow: Radio i svyaz'; 1989. 439 p. (in Russ.). ISBN 5-256-00180-9]
9. Farhang-Boroujeny B. *Adaptives Filters: Theory & Applications*. Wiley, UK; 1998. 529 p.
10. Shynk J.J. Frequency-domain and multirate adaptive filtering. *IEEE Signal Process. Mag*. 1992;9(1):14–37. https://doi.org/10.1109/79.109205
11. Popov D.I. Analysis of recursive rejection filters in transient mode. *Izvestiya Tul'skogo gosudarstvennogo universiteta. Tekhnicheskie nauki = Proceedings of Tula State University. Technical Sciences*. 2023;4:259–264. https://www.elibrary.ru/guzddw
12. Popov D.I. Optimization the parameters of recursive notch filters. *Izvestiya vysshikh uchebnykh zavedenii. Povolzhskii region. Tekhnicheskie nauki = University Proceedings. Volga Region. Engineering Sciences*. 2022;2(62):26–35. https://doi.org/10.21685/2072-3059-2022-2-2, https://www.elibrary.ru/fcnxab

13. Kulikov G.V., Konyashkin G.V. *Adaptive Notch Filter for Non-Fluctuation Interference Suppression:* RF Pat. 232764 U1. Publ. 19.03.2025.
14. Friedlander B., Porat B. The Modified Yule-Walker Method of ARMA Spectral Estimation. *IEEE Transactions on Aerospace Electronic Systems*. 1984;AES-20(2):158–173. https://doi.org/10.1109/TAES.1984.310437
15. Solonina A., Ulakhovich D. *Algoritmy i protsessory tsifrovoi obrabotki signalov* (*Algorithms and Processors of Digital Signal Processing*). St. Petersburg: BHV-Petersburg; 2002. 464 p. (in Russ.).
16. Nguyen Tien Phat. *Obrabotka radiotekhnicheskikh signalov na fone pomekh* (*Processing of Radio Signals Against the Background of Interference*). Monograph. Le Kui Don Vietnam State Technical University. Tambov: Konsaltingovaya kompaniya Yukom; 2018. 76 p. (in Russ.). https://www.elibrary.ru/xoyrql

СПИСОК ЛИТЕРАТУРЫ

1. *Помехозащищенность приема дискретных сигналов*; под ред. Ю.И. Савватеева, О.В. Назарова. М.: Радиотехника; 2015. 584 с. ISBN 978-5-93108-094-9
2. Борисов В.И., Зинчук В.М. *Помехозащищенность систем радиосвязи. Вероятностно-временной подход.* М.: РадиоСофт; 2008. 260 с. ISBN 5-93274-011-6. https://www.elibrary.ru/catzhm
3. Паршуткин А.В., Маслаков П.А. Помехоустойчивость каналов связи с амплитудно-фазовой модуляцией к воздействию непреднамеренных нестационарных помех. *Вопросы оборонной техники. Серия 16. Технические средства противодействия терроризму*. 2019;11–12:96–101.
4. Ложкин К.Ю., Петров А.В., Миронов В.А., Михалёв В.В., Прожеторко С.С. Аналитические зависимости средней вероятности искажения бита W-КАМ-сигнала на фоне гармонической или фазоманипулированной помех с учетом замираний. *Радиотехника*. 2020;84(4–8):27–35. https://doi.org/10.18127/j00338486-202004(8)-03
5. Куликов Г.В., Нестеров А.В., Лелюх А.А. Помехоустойчивость приема сигналов с квадратурной амплитудной манипуляцией в присутствии гармонической помехи. *Журнал радиоэлектроники.* 2018;11:2. https://doi.org/10.30898/1684-1719.2018.11.9
6. Куликов Г.В., Лелюх А.А., Баталов Е.В., Кузеленков П.И. Помехоустойчивость приема сигналов с квадратурной амплитудной манипуляцией в присутствии фазоманипулированной помехи. *Журнал радиоэлектроники.* 2019:7. URL: http://jre.cplire.ru/jre/jul19/10/text.pdf, https://doi.org/10.30898/1684-1719.2019.7.10
7. Куликов Г.В., Шамшура А.О., Печенин Е.А., Шаталов Е.В. Анализ помехоустойчивости приема сигналов с квадратурной амплитудной манипуляцией на фоне частотно-манипулированной помехи. *Вестник Воронежского института ФСИН России.* 2022;2:9–15.
8. Уидроу Б., Стирнз С. *Адаптивная обработка сигналов:* пер. с англ. М.: Радио и связь; 1989. 439 с. ISBN 5-256-00180-9
9. Farhang-Boroujeny B. *Adaptives Filters: Theory & Applications*. Wiley, UK; 1998. 529 p.
10. Shynk J.J. Frequency-domain and multirate adaptive filtering. *IEEE Signal Process. Mag*. 1992;9(1):14–37. https://doi.org/10.1109/79.109205
11. Попов Д.И. Анализ рекурсивных режекторных фильтров в переходном режиме. *Известия Тульского государственного университета. Технические науки*. 2023;4:259–264. https://www.elibrary.ru/guzddw
12. Попов Д.И. Оптимизация параметров рекурсивных режекторных фильтров. *Известия высших учебных заведений. Поволжский регион. Технические науки.* 2022;2(62):26–35. https://doi.org/10.21685/2072-3059-2022-2-2, https://www.elibrary.ru/fcnxab
13. Куликов Г.В., Коняшкин Г.В. *Адаптивный режекторный фильтр для подавления нефлуктуационных помех*: пат. 232764 U1 РФ. Заявка № 2024135726; заявл. 28.11.2024; опубл. 19.03.2025. Бюл. № 8.
14. Friedlander B., Porat B. The Modified Yule-Walker Method of ARMA Spectral Estimation. *IEEE Transactions on Aerospace Electronic Systems*. 1984;AES-20(2):158–173. https://doi.org/10.1109/TAES.1984.310437
15. Солонина А., Улахович Д. *Алгоритмы и процессоры цифровой обработки сигналов*. СПб.: БХВ-Петербург; 2002. 464 с.
16. Нгуен Тьен Фат. *Обработка радиотехнических сигналов на фоне помех:* Монография. Вьетнамский государственный технический университет им. Ле Куй Дона. Тамбов: ООО «Консалтинговая компания Юком»; 2018. 76 с. https://www.elibrary.ru/xoyrql

About the Authors

Georgy V. Konyashkin, Assistant, Department of Radio Electronic Systems and Complexes, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: konyashkin@mirea.ru. Scopus Author ID 58894490900, RSCI SPIN-code 9101-3255, https://orcid.org/0009-0003-1669-6604

Gennady V. Kulikov, Dr. Sci. (Eng.), Professor, Department of Radio Electronic Systems and Complexes, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: kulikov@mirea.ru. Scopus Author ID 36930533000, RSCI SPIN-code 2844-8073, http://orcid.org/0000-0001-7964-6653

Об авторах

Коняшкин Георгий Викторович, ассистент, кафедра радиоволновых процессов и технологий, Институт радиоэлектроники и информатики, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: konyashkin@mirea.ru. Scopus Author ID 58894490900, SPIN-код РИНЦ 9101-3255, https://orcid.org/0009-0003-1669-6604

Куликов Геннадий Валентинович, д.т.н., профессор, кафедра радиоэлектронных систем и комплексов, Институт радиоэлектроники и информатики, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: kulikov@mirea.ru. Scopus Author ID 36930533000, SPIN-код РИНЦ 2844-8073, http://orcid.org/0000-0001-7964-6653

Translated from Russian into English by K. Nazarov
Edited for English language and spelling by Thomas A. Beavitt

Modern radio engineering and telecommunication systems

Современные радиотехнические и телекоммуникационные системы

UDC 621.396.969
https://doi.org/10.32362/2500-316X-2025-13-5-75-86
EDN RTWBZR

RESEARCH ARTICLE

Optimization criterion for spacecraft observation planning algorithms

Alexander V. Ksendzuk [1, @], Ivan A. Kuznetsov [2]

[1] *MIREA – Russian Technological University, Moscow, 119454 Russia*
[2] *MAK Vympel, Moscow, 125480 Russia*
[@] *Corresponding author, e-mail: ks_alex@mail.ru*

• **Submitted:** 25.02.2025 • **Revised:** 31.03.2025 • **Accepted:** 25.07.2025

Abstract

Objectives. One of the critical tasks of space monitoring is the planning of observations due to the quality and amount of information obtained depending on how well the observation plan is developed. However, the selection of a method for planning spacecraft observations is hampered by a lack of unified criteria for comparing different planning algorithms. Therefore, the work sets out to develop planning quality criteria on the basis of physical observation principles based on radar, radiotechnical, and optical monitoring approaches in order to analytically determine their main parameters and check these parameters numerically.

Methods. The proposed quality criteria are deterministic, limited in energy by signal strength and observation time. The limiting values of the quality criteria for fixed observation time are analytically determined. In order to obtain the values of the quality criteria for four scheduling algorithms, a computational experiment is carried out.

Results. The proposed "weight–observation time" quality criterion is used to compare different observation planning algorithms that take into account spacecraft priority and total observation time. In order to account for the structure of the total observation time, the "weight–observation structure" criterion is introduced. It is analytically confirmed that the limited criteria values differ for different scheduling methods. The conducted numerical experiment is used to confirm the nature of the change of criteria for different planning methods and parameters included in the criteria.

Conclusions. The proposed observation planning quality criteria, which are based on the physical observation principles by radiotechnical and optical means, are used to numerically compare the results of spacecraft observation planning to take into account the priority of observation, as well as observation time and structure (how many and how long are the intervals into which the total observation time is divided). The possibility of using the proposed "weight–observation time" and "weight–observation structure" criteria to compare different planning algorithms is confirmed by computational experiment. Therefore, it is reasonable to use the proposed criteria for optimization of scheduling algorithms or their numerical comparison for different satellite observation conditions.

Keywords: observation planning, quality criterion, spacecraft, spacecraft monitoring, conflict observation

For citation: Ksendzuk A.V., Kuznetsov I.A. Optimization criterion for spacecraft observation planning algorithms. *Russian Technological Journal.* 2025;13(5):75–86. https://doi.org/10.32362/2500-316X-2025-13-5-75-86, https://www.elibrary.ru/RTWBZR

Financial disclosure: The authors have no financial or proprietary interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

К вопросу выбора критериев качества алгоритмов планирования наблюдений за космическими аппаратами

А.В. Ксендзук [1, @],
И.А. Кузнецов [2]

[1] *МИРЭА – Российский технологический университет, Москва, 119454 Россия*
[2] *Межгосударственная акционерная корпорация «Вымпел», Москва, 125480 Россия*
[@] *Автор для переписки, e-mail: ks_alex@mail.ru*

• **Поступила:** 25.02.2025 • **Доработана:** 31.03.2025 • **Принята к опубликованию:** 25.07.2025

Резюме

Цели. Одна из важнейших задач мониторинга космического пространства – это планирование наблюдений за космическими аппаратами. От того, насколько хорошо составлен план наблюдений, зависят качество и объем получаемой информации. В настоящее время существует множество различных методов планирования наблюдений за космическими аппаратами, однако единые критерии, которые позволяют сравнить различные алгоритмы планирования, отсутствуют. Цель работы – на основе физических принципов наблюдения радиолокационными, радиотехническими и оптическими средствами мониторинга разработать критерии качества планирования, определить их основные параметры аналитически и проверить численно.

Методы. Предложенные критерии качества – детерминированные, ограниченные по энергии мощностью сигнала и временем наблюдения. Аналитически определены предельные значения критериев качества для фиксированного времени наблюдения. В вычислительном эксперименте для 4 алгоритмов планирования получены значения критериев качества.

Результаты. Для сравнения различных алгоритмов планирования наблюдений, учитывающих приоритет космического аппарата и общее времени его наблюдения, предложен критерий качества «вес – время наблюдения». Для учета структуры общего времени наблюдения введен критерий «вес – структура наблюдения». Аналитически показано, что значения критериев ограничены, а также различаются для разных методов планирования. Выполнен численный эксперимент, который подтвердил характер изменения критериев для различных методов планирования и параметров, входящих в критерии.

Выводы. Предложенные критерии качества планирования наблюдений основаны на физических принципах наблюдения радиотехническими и оптическими средствами и позволяют численно сравнить результаты планирования наблюдений за космическими аппаратами с учетом приоритетности наблюдения, времени наблюдения и его структуры. Вычислительный эксперимент подтвердил возможность применения предложенных критериев «вес – время наблюдения» и «вес – структура наблюдения» для сравнения различных алгоритмов планирования. Предложенные критерии целесообразно использовать для оптимизации алгоритмов планирования или их численного сравнения для различных условий наблюдения за космическими аппаратами.

Ключевые слова: планирование наблюдений, критерий качества, космический аппарат, мониторинг космических аппаратов, конфликтное наблюдение

Для цитирования: Ксендзук А.В., Кузнецов И.А. К вопросу выбора критериев качества алгоритмов планирования наблюдений за космическими аппаратами. *Russian Technological Journal*. 2025;13(5):75–86. https://doi.org/10.32362/2500-316X-2025-13-5-75-86, https://www.elibrary.ru/RTWBZR

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

At present, a significant increase in the number of spacecraft (SC) and quantity of space debris is accompanied by a relatively slow growth in the quantity and quality of optical, radar and radiotechnical systems of monitoring near-Earth space[1] [1]. Under these conditions, the task of planning observations by monitoring means acquires special importance. The relevance of this problem is confirmed by the works [2–8], in which the authors propose various algorithms for planning observations, including those of astronomical objects [9, 10].

In order to evaluate the effectiveness of the scheduling algorithm, the influence of each factor (weather conditions, particle dispersion level, etc.) on the final result is distributed in a percentage based on the influence of each factor on the quality of the problem solution [2]. In this work, each parameter is evaluated in the range from 1 to 10 points, then the value of the parameter is multiplied by its corresponding importance percentage, and all weighted scores are summarized.

In [3], maps of SC detection efficiency are constructed for observation planning purposes using the resulting efficiency factor, which is calculated as the product of criteria (extinction coefficient, angular velocity, etc.) together with its weighting factor.

The use of a covariance matrix trace for describing the average change in the contribution to the measurements when selecting the SC to be monitored is set out in [4]. Such a trace can be related to the informativity parameters, for example, the change of differential entropy.

In [5], the effectiveness of "greedy" optimization methods that do not require significant computational and mathematical resources is analyzed. As a criterion, a cost function is used based on potential observations and how well the observations fit the objective, while numerical and analytical formulas for calculating "fit" are not given.

In [6], optimization is performed on the object observation time, taking into account the observation switching time from the previous object. Although the described global optimization approach can be used to maximize the total SC observation time, the authors do not take priority of SC observation into account.

The authors of [7] describe the use of average satellite observation time minus the deviation of the observation time of each satellite with respect to the mean value multiplied by the Lagrange factor to obtain an observation criterion. This criterion takes the maximum value when all satellites are observed for the same time.

Thus, despite the availability of a large number of methods, algorithms, and scheduling software, there are no generally accepted quality criteria that allow us to compare these methods. Consequently, the development of physically based criteria becomes an urgent task.

1. PLANNING TASK STATEMENT

When planning SC observations, one of the main problems concerns how to resolve conflicts when the number (bandwidth) of observation channels is less than the number of simultaneously visible SC, i.e., the need to choose which SC should be observed [11]. SC visibility is understood as the possibility of its observation in space (geometric visibility), in energy parameters (radar and radiotechnical visibility) and in the frequency domain (for radiotechnical systems, the frequency of the SC signal must be within the range of operating frequencies).

If there were no conflicts, the task of forming an observation plan would be reduced to sequential tracking of visible SCs and, if necessary to save the resource of observation facilities, stopping observations when the required amount of information has been obtained.

In realistic observing conditions involving a lack of observational means, conflicts frequently arise in terms of the simultaneous visibility of multiple satellites, for each of which a decision must be made as to which of the SCs to observe. Different conflict resolution methods lead to the formation of different observation plans.

The initial data for planning are: the matrix of SC visibility, SC priority and time sufficient for obtaining information of the required quality. For the convenience of planning, the time is assumed to be discrete.

In the **visibility matrix V**(i, t), the row i corresponds to the SC number, while column t corresponds to the time, and each cell $V_{i,\,t}$ shows the visibility of the SC

[1] https://www.space-track.org/. Accessed January 20, 2025.

by the monitoring tool. In the simplest case, the values in the matrix take the values 0 and 1 (visibility based on geometric relations). In a more complex variant, the cells contain a value related to the quality of observation, for example, the signal-to-noise ratio by power or the probability of SC observation. The sum of the rows within a column indicates how many SCs are visible (can be observed) at the same time.

The vector of priority (weights) of the SC for each SC defines a weight w_i representing a positive integer that indicates the importance of the SC observation (the larger the value, the higher the priority and value of the observation of this SC).

The vector of sufficient (continuous) observation time for each *i*th SC determines the (continuous) time interval $T_{\text{suf}\,i}$, sufficient to obtain the required amount of information/quality of parameter estimates. The requirement of interval continuity is based on the principles for estimating parameters constantly over the observation interval. If this requirement is optional, $T_{\text{suf}\,i}$ can be the full observation time. In cases when the signal-to-noise power ratio is used in the observation matrix, $T_{\text{suf}\,i}$ defines the power signal-to-noise ratio on which the quality of estimates depends [12].

The result of planning is the **observation matrix O** (i, t), the dimensionality of which coincides with the dimensionality of the visibility matrix. For each *i*th row (SC number) at time *t* in the cell $O_{i,t}$ can be the value 1 (*i*th SC at this time will be observed by the monitoring instrument) or 0 (SC will not be observed). At one moment of time *t* (column of the matrix), the number of SC being monitored shall be equal to the number of channels of the monitoring instrument.

An example showing the type of input data for planning and the result of planning is shown in Fig. 1.

2. NUMERICAL OBSERVATION RATES

Numerical parameters used to calculate the quality criterion and analyze the observation and visibility matrices are described as follows:

$t = 1 \ldots T_{\text{plan}}$, T_{plan} is the planning interval (time interval for which the observation plan is prepared, usually a day).

$i = 1 \ldots I_{\max}$ is the number of visible SCs in the planning interval T_{plan}.

$T_{\text{vis}\,i}$ is the visibility time of the *i*th SC is the sum of its free (conflict-free) $T_{\text{vis.free}\,i}$ and conflict $T_{\text{vis.conf}\,i}$ visibility time $T_{\text{vis}\,i} = T_{\text{vis.free}\,i} + T_{\text{vis.conf}\,i} = \Sigma_{t=1\ldots T} V_{i,t}$. The visibility time shows the potential observation capability of the SC.

$I_{\text{obs}} \le I_{\max}$ is the number of observed SCs in the planning interval T_{plan}.

$T_{\text{obs}\,i}$ is the observation time for the *i*th SC is the sum of the conflict-free $T_{\text{obs.free}\,i}$ and conflict $T_{\text{obs.conf}\,i}$ observation time. These values show the realization of a particular observation plan, with $T_{\text{obs}\,i} \le T_{\text{vis}\,i}$.

w_i is the priority (weight) of the SC, which indicates the value (contribution) of its observation. The priority is set by an integer number lying in a given range $\{w \in \mathbb{N} \mid 1 \le w_i \le w_{\max}\}$; the larger it is, the higher the priority of the SC (in most real-world tasks, it is sufficient to set w_i values from 1 to 5; this range is used in numerical calculations in this study).

$T_{\text{suf}\,i}$ is the sufficient (continuous) time of SC observation. It is determined for each satellite depending on the observation task (signal detection, estimation of signal parameters, estimation of parameters of the transmitted information flow, etc.), known, predicted or calculated statistical characteristics of the observation (signal-to-noise ratio, probability of satellite observation). Observation of the SC during the (continuous) time

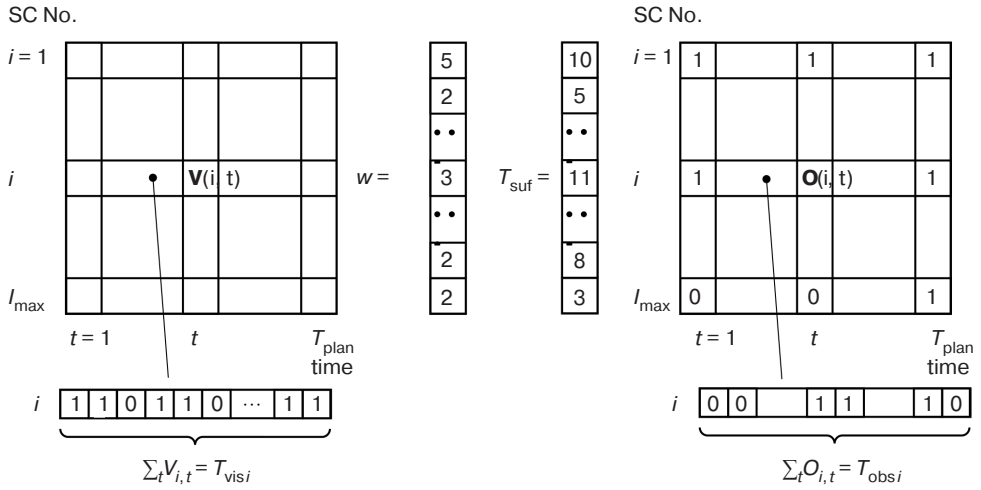


Fig. 1. Type of basic matrices and vectors when planning observations

interval $T_{\text{suf}\,i}$ provides with the required probability the necessary quality of its parameter estimation. Observation for a shorter time interval does not provide the required quality. Observation for a longer interval is excessive.

$T_{\text{vis.s}}$ is the total (sum) time of visibility (at least one of $i = 1 \ldots I_{\max}$ SCs can be observed), determined by the visibility matrix $\mathbf{V}(i, t)$:

$$T_{\text{vis.s}} = \sum_{t=1\ldots T} T_{\text{vis.s}\,t} = \sum_{t=1\ldots T} \text{sign}\left[\sum_{i=1\ldots I_{\max}} V_{i,t}\right]. \quad (1)$$

N_t is the number of simultaneously visible SCs at the moment of time t:

$$N_t = \left[\sum_{i=1\ldots I_{\max}} V_{i,t}\right]. \quad (2)$$

$T_{\text{obs.s}}$ is the total observation time (at least one of $i = 1 \ldots I_{\max}$ SCs), is determined by the observation matrix $\mathbf{O}(i, t)$:

$$T_{\text{obs.s}} = \sum_{t=1\ldots T} T_{\text{obs.s}\,t} = \sum_{t=1\ldots T} \text{sign}\left[\sum_{i} O_{i,t}\right]. \quad (3)$$

While the free and conflict observation time can be integrally determined by analyzing the visibility matrix $\mathbf{V}(i, t)$, the calculation of the redistribution of this time between the satellites is based on an analysis of the corresponding observation matrix $\mathbf{O}(i, t)$ for this planning method. As an example, Fig. 2 shows the result of scheduling observations by a single-channel SC constellation tool in the presence of free and conflict observations.

The observation matrix depicted in Fig. 2 confirms the conservation of free observation time for all SC. **The free observation time** is preserved unless **the resource limitation of the observation time** $T_{\text{obs.s}} \le T_{\text{res}}$, $t = 1 \ldots T_{\text{plan}}$ is used to extend the lifetime of the monitoring instruments or to use part of the time for other purposes, e.g., to analyze the environment (interference emissions are evaluated using radiotechnical and radar instruments, while astroclimatic factors are examined by means of optical monitoring instruments). In the absence of such restrictions, one of the methods for checking the correctness of the planning algorithm $T_{\text{vis.free}\,i} = T_{\text{obs.free}\,i}$ is preserving the time of free (conflict-free) observation.

Conflict observation time is distributed among the SCs according to the planning method used (including the simplest method involving termination of observations in case of conflict). In the example considered in Fig. 2, the priority is given to the SCs at the beginning

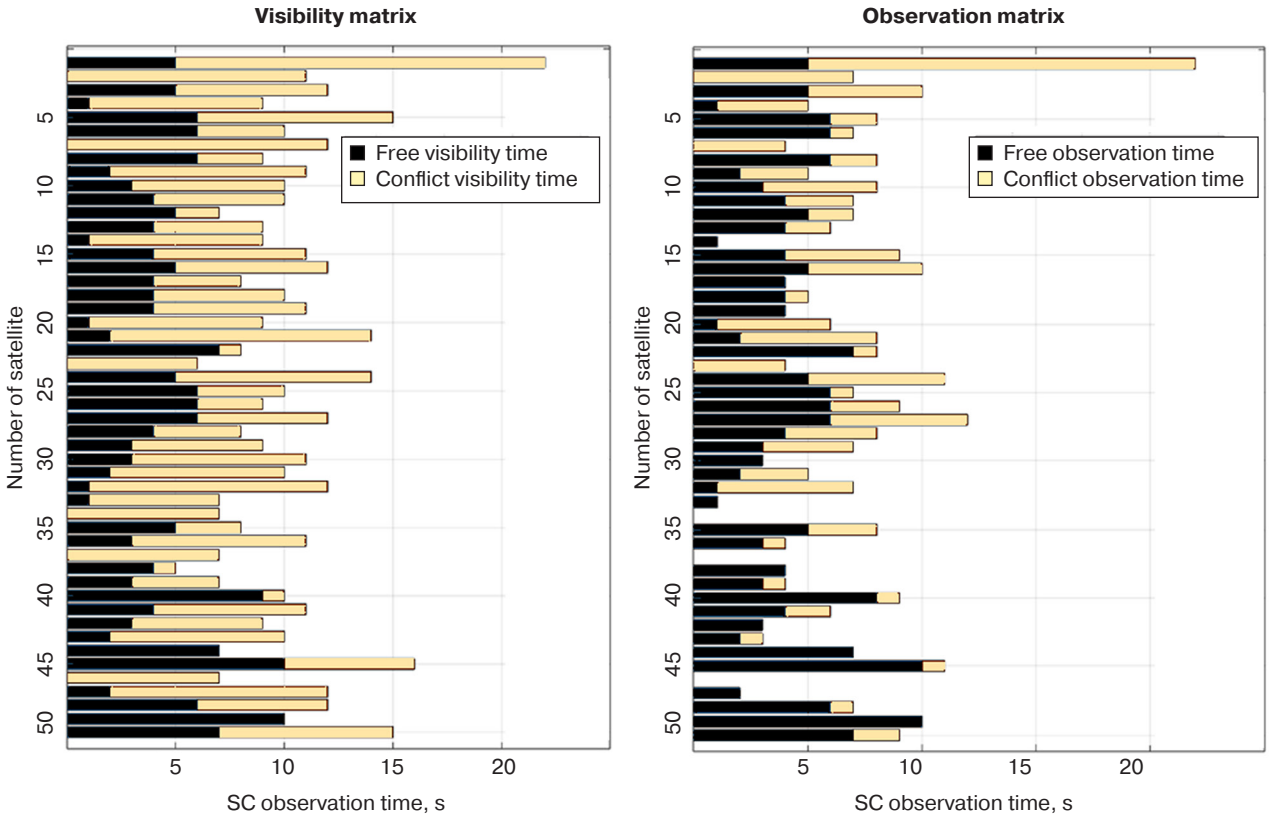


Fig. 2. Example of the visibility matrix (left) and observation plan (right)

of the list (all possible conflict observation time is used), while for the rest the time is redistributed.

Total observation time in the absence of resource constraint should coincide with the total visibility time $T_{\text{vis.s}} = T_{\text{obs.s}}$, which should be used to verify the correctness of the scheduling algorithm.

3. QUALITY CRITERIA FOR OBSERVATIONS

To compare different planning methods, it is necessary to introduce a numerical indicator based on numerical, measurable parameters of the observation plan [13]. Since in the considered case visibility and planning result are deterministic functions, we will determine the quality of planning without using probabilistic indicators. However, it is not difficult to take probabilistic parameters into account within the framework of the proposed approach—it is enough to put the probabilities of SC observation pre-calculated or determined on the basis of statistical data as the values of elements in the cells [14].

3.1. "Weight–observation time" criterion

The first planning quality criterion proposed in this paper is introduced using the priority of the SC and the total time of its observation (without taking into account how many intervals it is divided into). This variant can be used for the case when one discrete time interval is sufficient to estimate the parameters of the SC with the required quality. The criterion "weight–observation time" is defined by the sum of **the contributions from the observation of the individual SCs** Q_{wt_i}:

$$Q_{wt} = \frac{1}{I_{\max}} \sum_{i=1}^{I_{\max}} Q_{wt_i} = \frac{1}{I_{\max}} \sum_{i=1}^{I_{\max}} (a_i)^n \ln\left(\frac{T_{\text{obs}\,i}}{T_{\text{suf}\,i}} + 1\right), \quad (4)$$

where:

- $a_i = w_i/w_{\max}$ is the relative priority of the ith SC, taking values in the interval (0; 1];
- w_i is the priority of the ith SC;
- $w_{\max}$ is the maximum observation priority from the set of visible SCs;
- n is the relative priority degree indicator, making it possible to vary the contribution of w_i: $n = 1$ is the linear dependence on the priority of the SC, $n = 0$ is the priority of the SC is not taken into account;
- $I_{\max}$ is the amount of the visible SCs;
- $T_{\text{obs}\,i}$ is the observation time of the ith SC obtained as a result of planning;
- $T_{\text{suf}\,i}$ is the time sufficient to obtain information / estimates of parameters of the required quality for the ith SC.

The proposed criterion "weight–observation time" has the following properties:

1. If the ith SC is not observed (i.e., $T_{\text{obs}\,i} = 0$) in the planning interval T_{plan}, but is included in the set of visible satellites $I_{\max}$, the contribution from this SC is equal to zero:

$$Q_{wt_i}\big|_{T_{\text{obs}\,i=0}} = \left(\frac{w_i}{w_{\max}}\right)^n \ln(1) = 0. \quad (5)$$

2. If the observation time of the ith SC is equal to sufficient time ($T_{\text{obs}\,i} = T_{\text{suf}\,i}$), the contribution from its observation will depend on the SC priority:

$$Q_{wt_i} \approx 0.69\left(\frac{w_i}{w_{\max}}\right)^n, \quad (6)$$

while for the SC with the maximum priority, the contribution will be equal to 0.69. The same value of Q_{wt_i} will be for a satellite with any priority at $n = 0$ and observation for sufficient time ($T_{\text{obs}\,i} = T_{\text{suf}\,i}$).

3. The contribution from the observation of the ith SC has a maximum (bounded from above) value when the time of SC observation tends to the time of its visibility $T_{\text{obs}\,i} \to T_{\text{vis}\,i}$. The limit for continuously visible SCs is equal to the planning time $T_{\text{obs}\,i} \to T_{\text{plan}}$:

$$\max(Q_{wt_i}) = \left(\frac{w_i}{w_{\max}}\right)^n \ln\left(\frac{T_{\text{vis}\,i}}{T_{\text{suf}\,i}} + 1\right). \quad (7)$$

For the SC with the maximum priority (or for any SC if the observation priority is not taken into account), $n = 0$) $\max(Q_{wt_i}) = \ln\left(\frac{T_{\text{vis}\,i}}{T_{\text{suf}\,i}} + 1\right)$.

A diagram of the SC observation contribution Q_{wt_i} as a function of the observation time $T_{\text{obs}\,i}$ versus the sufficient observation time $T_{\text{suf}\,i}$ at different priority ($w_i = 1 \ldots 5$) and its degree of consideration $n = 0, 0.5, 1, 2$ is shown in Fig. 3.

Analyzing the results of numerical calculation of the proposed "weight–observation time" criterion confirms the above properties:

- Q_{wt_i} monotonically increases with increasing observation time;
- at small observation time relatively sufficient $T_{\text{obs}\,i}/T_{\text{suf}\,i} \to 0$ the criterion "weight–observation time" $Q_{wt_i} \to 0$;
- for the SC with the maximum priority $w_i = 5$ the contribution, depending on the relative observation time $T_{\text{obs}\,i}/T_{\text{suf}\,i}$ grows logarithmically, taking at

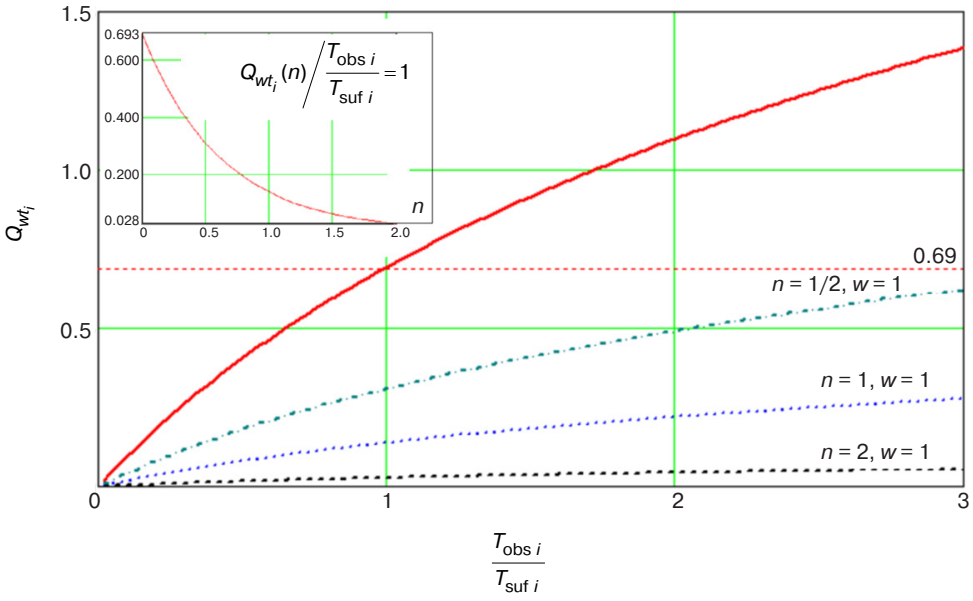


Fig. 3. The value of the quality index Q_{wt_i} at satellite observation depending on the parameters (n, w) and the value of Q_{wt_i} (n; $w = 1$) at $T_{obs\ i}/T_{suf\ i} = 1$

the point where the observation time of the ith SC is equal to a sufficient $T_{obs\ i} = T_{suf\ i}$, value of 0.69. The value of n does not affect the position of this graph, since $(a_i)^n = \left(\frac{w_{max}}{w_{max}}\right)^n = 1$;

- for the SC with minimum priority ($w_i = 1$, $w_{max} = 5$), the contribution for the same relative observation time $T_{obs\ i}/T_{suf\ i}$ is less and is determined by the degree n, which takes into account the weight of the SC priority: the larger is n, the less is the contribution from the observation of SC with low priority. In the considered example, at the point $T_{obs\ i} = T_{suf\ i}$ for $n = [0 \ldots 2]$ $Q_{wt_i} = [0.69 \ldots 0.028]$, the diagram is shown in the inset of Fig. 3.

3.2. "Weight–observation structure" criterion

For SCs whose signal parameters may change over time, continuous long observation intervals are of great value (e.g., it is more accurate to determine a linearly time-varying frequency from a single long-time interval rather than from a set of short ones).

In order to take the structure of the observation time into account, i.e., how many intervals of what duration into which the total observation time of the SC $T_{obs\ I}$ is divided, it is proposed to introduce the equivalent observation time $T_{obs.eq\ i}$, which is calculated by the formula:

$$T_{obs.eq\ i} = \frac{\sum_{k=1}^{K_i} \left(T_{obs\ i_k}\right)^2}{T_{obs\ i}}, \qquad (8)$$

where K_i is the number of noncontiguous observation intervals of the ith SC during the period T_{plan}; k is the number of the observation site of the ith SC; $T_{obs\ i_k}$ is the observation time at the kth site.

Figure 4 shows the dependence of the equivalent observation time $T_{obs.eq}$ of a SC on the continuous interval $T_{obs\ k}$ of its observation reduced to the total observation time T_{obs} of the SC.

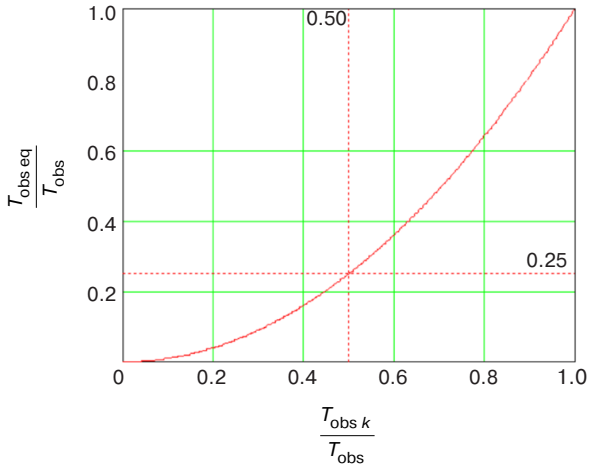


Fig. 4. Dependence of $T_{obs\ eq}$ values on the observation interval $T_{obs\ k}$ when normalized by the total satellite observation time T_{obs}

The properties of the equivalent observation time following from expression (8) and confirmed by the diagram in Fig. 5:

1. For one SC observation interval ($K = 1$), the equivalent observation time is equal to the observation time $T_{obs\ i}$ itself:

$$T_{\text{obs.eq } i} = \frac{\left(T_{\text{obs } i_k}\right)^2}{T_{\text{obs } i}} = T_{\text{obs } i}, \qquad (9)$$

2. At a large number of observation time intervals ($K_i \gg 1$) and one "long" interval $T_{\text{obs } i_k} \gg T_{\text{obs } i_j}$ the equivalent observation time of the SC is determined by the maximum duration of the site $\max(T_{\text{obs} i_k})$.
3. When the number of sites $K_i \to \infty$ with the same small observation time $T_{\text{obs } i_k}$ is large, the equivalent observation time tends to zero:

$$T_{\text{obs.eq } i} = \frac{\sum_{k=1}^{K_i \to \infty} \left(T_{\text{obs } i_k}\right)^2}{T_{\text{obs } i}} \to 0. \qquad (10)$$

The second "weight–observation structure" planning quality criterion proposed in this article uses in (4), instead of the total observation time of the ith SC $T_{\text{obs } i}$ its equivalent observation time $T_{\text{obs.eq } i}$:

$$Q_{ws} = \frac{1}{I_{\max}} \sum_{i=1}^{I_{\max}} Q_{ws_i} = \frac{1}{I_{\max}} \sum_{i=1}^{I_{\max}} (a_i)^n \ln\left(\frac{T_{\text{obs eq}\, i}}{T_{\text{suf } i}} + 1\right) =$$

$$= \frac{1}{I_{\max}} \sum_{i=1}^{I_{\max}} \left(\frac{w_i}{w_{\max}}\right)^n \ln\left(\frac{\sum_{k=1}^{K} \left(T_{\text{obs } i_k}\right)^2}{T_{\text{obs } i} T_{\text{suf } i}} + 1\right), \qquad (11)$$

where the relative priority of the ith SC $a_i^n = (w_i / w_{\max})^n$ and the ratio of the equivalent observation time of the ith spacecraft $T_{\text{obs.eq } i}$ to the continuous interval $T_{\text{suf } i}$ sufficient for obtaining estimates of the required quality are taken into account.

The "weight–observation structure" criterion Q_{ws} has the same properties as the "weight–observation time" criterion Q_{wt} with the precision that the observation time should be replaced by an equivalent one.

4. ASSESSMENT OF THE APPLICABILITY OF THE PLANNING QUALITY CRITERIA

In order to evaluate the feasibility of the proposed quality criteria, we use several different scheduling algorithms for which the behavior of the proposed criteria Q_{wt} (4) and Q_{ws} (11) can be logically predicted. For these scheduling algorithms, we use the same visibility matrix for which we compute the observation matrix and the proposed quality criteria. The results of the behavior of the criteria, including their compliance with the logical assumptions and sufficient numerical spread for different planning algorithms, will confirm the possibility of applying the proposed quality criteria.

Algorithm 1 ("first on the list")

In case of a conflict observation, an unconditional transition to the observation of the SC that is first in the list is performed. Obviously, such an algorithm will give low time (continuous) observation of the SC at the end of the list. If the SCs with high priority are located there, the quality of planning by the Q_{wt} criterion will be low. The value of the criterion will significantly depend on which SC with what priority are located first in the list.

In order for such a simple and computationally undemanding scheduling algorithms to be practically useful, it is necessary to compile a list of the SCs, placing the highest-priority ones with the minimum total observation time at the top of the list.

Algorithm 2 ("inertia-free transition to priority satellite")

In case of a conflicting observation (collision), an unconditional switch to the observation of the SC that has a higher priority regardless of the time of its observation. In fact, this is a "greedy" algorithm that maximizes the quality criterion at each step [15]. Obviously, such an algorithm will give a higher (compared to Algorithm 1) quality of observation for criterion (4) Q_{wt} "weight–observation time"; for the criterion Q_{ws}, the contribution of lower-priority SCs will be low, since high-priority SCs will be observed in conflict observation. If the priority of all SCs is equal, the criterion Q_{ws} (11) "weight–continuous observation time" will be the same as for Algorithm 1.

Algorithm 3 ("inertial transition to a satellite with higher weight")

A modification of Algorithm 2 takes into account the observation time of the ith SC by calculating a coefficient equal to the ratio of the priority of the SC to the time it has already been observed ($w_i/T_{\text{obs } i}$). In contrast to Algorithm 2, the transition to the observation of a new, higher priority jth SC may not occur if the time $T_{\text{obs } j} \gg T_{\text{obs } i}$. That is, the formed "inertia" of the SC observation decreases when its observation time increases. Such a criterion should lead to a decrease in short observation intervals due to inertia and to an increase in the quality criterion (6) compared to Algorithms 1 and 2.

Algorithm 4 ("observation time equalization")

In this case, tracking of a SC is stopped if its observation time exceeds the same parameter of another satellite. The algorithm tries to "equalize" the observation time of all SCs.

To determine the maximum achievable indicators on the visibility matrix, it is reasonable to use the method of complete enumeration over all possible variants of the planning problem solution. However, for real conditions (plan for a day, more than 10000 SC overflights) such a solution can be theoretically obtained only with the use of quantum computing [16]. However, such a solution can also be obtained by restricting the list of satellites to those observed by radiotechnical systems of RTU MIREA [17].

An illustration of the working principle of Algorithms 1–4, for which the planning quality criterion is calculated, is shown in Fig. 5.

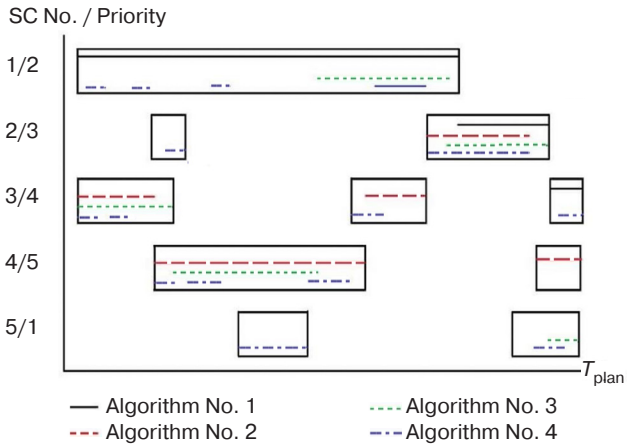


Fig. 5. Example of operation of Algorithms 1–4 on planning interval T_{plan} for 5 SCs with given visibility time and different observability priority a_i

Table summarizes the results of the calculation according to the proposed quality criteria. As input data, a SC visibility matrix consisting of 50 SC for the discrete planning interval $T_{plan} = 500$ was generated. For the table columns, the minimum values of the quality criteria are italicized, while the maximum values are given in bold.

Table. Results of numerical modeling

Algorithm designation	Q_{wt}		Q_{ws}	
	$n = 0$	$n = 1$	$n = 0$	$n = 1$
Algorithm 1	*0.382*	*0.255*	*0.219*	*0.148*
Algorithm 2	0.576	0.404	0.321	0.210
Algorithm 3	**0.618**	**0.530**	**0.370**	**0.295**
Algorithm 4	0.480	0.309	0.258	0.189

The numerical modeling results confirm the analytical conclusions and logical assumptions as follows:

- criteria Q_{wt} and Q_{ws} take maximum values for Algorithm 3 both when taking into account $n = 1$ and without taking into account $n = 0$ priority of the SC, because by its principle it has inertia, leading to the formation of longer observation sites, which is confirmed by Fig. 5;
- criteria Q_{wt} and Q_{ws} take minimum values for Algorithm 1 both when taking into account $n = 1$ and without taking into account $n = 0$ priority of the SC by enabling the observation of satellites at the beginning of the list without taking into account their priority (in the considered case, it is low) as confirmed by Fig. 5;
- the ratio of the minimum and maximum values of the criteria Q_{wt} is 2.078 when taking into account the priority of SC $n = 1$ and 1.617 without taking into account the priority $n = 0$, respectively. Consequently, the difference between the maximum and minimum values of the criterion "weight–observation time" is increased by taking into account the priority;
- the ratio of the minimum and maximum values of the criteria Q_{ws} is 1.993 when taking into account the priority of the SC and 1.689 without taking into account the priority. Similarly to the Q_{wt} criterion, taking into account the priority increases the difference between the maximum and minimum values of the criterion "weight–structure of observation".

CONCLUSIONS

Under objective conditions of satellite orbital launch dynamics in relation to the increase in the number and capabilities of their monitoring means, the task of observation planning is of particular importance. Two quality criteria are proposed for comparing different methods of SC observation planning: "weight–observation time" and "weight–observation structure". The criteria are based on the formation of a value associated with the energy signal-to-noise ratio, which determines the quality of parameter estimates, as well as taking into account the priority of satellites and the structure of the total observation time (number of intervals).

Four scheduling algorithms used for verification give predictable results for the relative scheduling quality according to the introduced criteria. Numerical calculations of the quality criteria confirmed the theoretical assumptions on which they are based. For the same SC observation conditions, the quality criteria of the plan obtained by four different scheduling algorithms differ more than 1.5 times.

Further development of criteria is envisaged in terms of their specification, for example, for radiotechnical systems of monitoring available in "Cosmocenter" at the RTU MIREA synthesis of optimal methods of planning observations and their experimental confirmation.

Authors' contributions

A.V. Ksendzuk—formulating the problem; developing the criteria, key relationships, and conclusions.

I.A. Kuznetsov—developing the software for verification of calculations; analysis of simulation results.

REFERENCES

1. Liu J., Yang X., Cheng H., et al. Progress of China's Space Debris Research. *Chinese J. Space Sci.* 2022;42(4):824–829. https://doi.org/10.11728/cjss2022.04.yg26
2. Cowles K. Site selection criteria for the optical atmospheric visibility monitoring telescopes. *The Telecommunications and Data Acquisition Report* (*TDA Progress Rep*.). 1989;42–99:235–239.
3. Elenin L.V., Molotov I.E., Borovin G.K. Effective planning of observations of space objects on different types of orbits. *Preprinty Instituta prikladnoi matematiki im. M.V. Keldysha RAN = Preprints of the Keldysh Institute of Applied Mathematics.* 2018;72. 18 p. (in Russ.). https://doi.org/10.20948/prepr-2018-72
4. Fedeler S.J., Holzinger M.J., Whitacre W. Optimality and Application of Tree Search Methods for POMDP-based Sensor Tasking. In: *Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference* (*AMOS*). Maui, Hawaii, USA. 2020. 24 p. Available from URL: https://amostech.com/TechnicalPapers/2020/Poster/Fedeler.pdf
5. Schubert M., Kebschull C., Gelhaus J., et al. Evaluating sensor tasking strategies for object cataloging in GEO. *Acta Astronautica.* 2024;228:7–16. https://doi.org/10.1016/j.actaastro.2024.11.026
6. Dhingra N.K., DeJac C., Herz A., et al. Space domain awareness sensor scheduling with optimality certificates. In: *Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference*. 2023. 15 p. Available from URL: https://amostech.com/TechnicalPapers/2023/SDA/Dhingra.pdf
7. Ksendzuk A., Grigorev V. Satellite Radio Monitoring Stations Observation Planning: Time Alignment Observation Algorithm. In: *2021 International Conference Engineering and Telecommunication* (*En&T*). IEEE; 2021. P. 2–6. https://doi.org/10.1109/EnT50460.2021.9681763
8. Tian M., Ma G., Huang P., et al. Optimizing satellite ground station facilities scheduling for RSGS: a novel model and algorithm. *Int. J. Digital Earth.* 2023;16(1):3949–3972. https://doi.org/10.1080/17538947.2023.2259870
9. Garcia-Piquer A., Morales J.C., Ribas I., et al. Efficient scheduling of astronomical observations –Application to the CARMENES radial-velocity survey. *Astronomy & Astrophysics.* 2017;604:A87. https://doi.org/10.1051/0004-6361/201628577
10. Johnston M.D. Scheduling tools for astronomical observations. In: Boroson T.A., Davies J.K., Robson I. (Eds.). *New Observing Modes for the Next Century. Astronomical Society of the Pacific Conference Series* (*ASP*). 1996;87:62–71. Available from URL: https://ui.adsabs.harvard.edu/link_gateway/1996ASPC...87...62J/ADS_PDF
11. Grigorev V.S., Ksendzuk A.V. Optimization methods for scheduling observations of spacecraft by ground-based radio measuring instrument. *Zhurnal Radioelektroniki = J. Radio Electronics.* 2023;7 (in Russ.). https://doi.org/10.30898/1684-1719.2023.7.1
12. Dulevich V.E. (Ed.). *Teoreticheskie osnovy radiolokatsii* (*Theoretical Foundations of Radar*). Moscow: Sovetskoe Radio; 1978. 607 p. (in Russ.).
13. Fal'kovich S.E., Khomyakov E.N. *Statisticheskaya teoriya izmeritel'nykh radiosistem* (*Statistical Theory of Measurement Radio Systems*). Moscow: Radio i svyaz'; 1981. 965 p. (in Russ.).
14. Fürbacher A., Fruth T., Weibigke A., et al. Concept for generic agile, reactive optical link planning. *CEAS Space J.* 2025. 10 p. https://doi.org/10.1007/s12567-025-00592-0
15. García A. Greedy algorithms: a review and open problems. *ArXiv Prepr.* arXiv:2408.08935 (2024). https://doi.org/10.48550/arXiv.2408.08935
16. Sigov A.S., Andrianova E.G., Zhukov D.O., Zykov S.V., Tarasov I.E. Quantum informatics: Overview of the main achievements. *Russian Technological Journal*. 2019;7(1):5–37 (in Russ.). https://doi.org/10.32362/2500-316X-2019-7-1-5-37
17. Ksendzuk A.V., Zamuruev S.N. Prospects for the creation of a radio engineering complex for monitoring outer space on the basis of the MIREA Space Center. In: *Actual Problems and Prospects of Development of Radio Engineering and Information Communication Systems* (*RADIOINFOCOM-2022*): *Proceedings of the 6th Scientific and Technical Committee*. Moscow: RTU MIREA; 2022. P. 72–75 (in Russ.). https://www.elibrary.ru/knywzk

СПИСОК ЛИТЕРАТУРЫ

1. Liu J., Yang X., Cheng H., et al. Progress of China's Space Debris Research. *Chinese J. Space Sci.* 2022;42(4):824–829. https://doi.org/10.11728/cjss2022.04.yg26
2. Cowles K. Site selection criteria for the optical atmospheric visibility monitoring telescopes. *The Telecommunications and Data Acquisition Report* (*TDA Progress Rep*.). 1989;42–99:235–239.

3. Еленин Л.В., Молотов И.Е., Боровин Г.К. Эффективное планирование наблюдений космических объектов на орбитах различных типов. *Препринты Института прикладной математики им. М.В. Келдыша РАН.* 2018;72. 18 с. https://doi.org/10.20948/prepr-2018-72
4. Fedeler S.J., Holzinger M.J., Whitacre W. Optimality and Application of Tree Search Methods for POMDP-based Sensor Tasking. In: *Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference* (*AMOS*). Maui, Hawaii, USA. 2020. 24 p. URL: https://amostech.com/TechnicalPapers/2020/Poster/Fedeler.pdf
5. Schubert M., Kebschull C., Gelhaus J., et al. Evaluating sensor tasking strategies for object cataloging in GEO. *Acta Astronautica.* 2024;228:7–16. https://doi.org/10.1016/j.actaastro.2024.11.026
6. Dhingra N.K., DeJac C., Herz A., et al. Space domain awareness sensor scheduling with optimality certificates. In: *Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference*. 2023. 15 p. URL: https://amostech.com/TechnicalPapers/2023/SDA/Dhingra.pdf
7. Ksendzuk A., Grigorev V. Satellite Radio Monitoring Stations Observation Planning: Time Alignment Observation Algorithm. In: *2021 International Conference Engineering and Telecommunication* (*En&T*). IEEE; 2021. P. 2–6. https://doi.org/10.1109/EnT50460.2021.9681763
8. Tian M., Ma G., Huang P., et al. Optimizing satellite ground station facilities scheduling for RSGS: a novel model and algorithm. *Int. J. Digital Earth.* 2023;16(1):3949–3972. https://doi.org/10.1080/17538947.2023.2259870
9. Garcia-Piquer A., Morales J.C., Ribas I., et al. Efficient scheduling of astronomical observations –Application to the CARMENES radial-velocity survey. *Astronomy & Astrophysics.* 2017;604:A87. https://doi.org/10.1051/0004-6361/201628577
10. Johnston M.D. Scheduling tools for astronomical observations. In: Boroson T.A., Davies J.K., Robson I. (Eds.). *New Observing Modes for the Next Century. Astronomical Society of the Pacific Conference Series* (*ASP*), 1996;87:62–71. URL: https://ui.adsabs.harvard.edu/link_gateway/1996ASPC...87...62J/ADS_PDF
11. Григорьев В.С., Ксендзук А.В. Оптимизационные методы составления расписания наблюдений за космическими аппаратами в околоземном пространстве наземными радиотехническими измерительными средствами. *Журнал радиоэлектроники.* 2023;7. https://doi.org/10.30898/1684-1719.2023.7.1
12. *Теоретические основы радиолокации*; под ред. В.Е. Дулевича. М.: Сов. радио; 1978. 607 с.
13. Фалькович С.Е., Хомяков Э.Н. *Статистическая теория измерительных радиосистем.* М.: Радио и связь; 1981. 965 с.
14. Fürbacher A., Fruth T., Weibigke A., et al. Concept for generic agile, reactive optical link planning. *CEAS Space J.* 2025. 10 p. https://doi.org/10.1007/s12567-025-00592-0
15. García A. Greedy algorithms: a review and open problems. *ArXiv Prepr.* arXiv:2408.08935 (2024). https://doi.org/10.48550/arXiv.2408.08935
16. Сигов А.С., Андрианова Е.Г., Жуков Д.О., Зыков С.В., Тарасов И.Е. Квантовая информатика: обзор основных достижений. *Russian Technological Journal*. 2019;7(1):5–37. https://doi.org/10.32362/2500-316X-2019-7-1-5-37
17. Ксендзук А.В., Замуруев С.Н. Перспективы создания радиотехнического комплекса мониторинга космического пространства на базе космоцентра МИРЭА. В сб.: *Актуальные проблемы и перспективы развития радиотехнических и инфокоммуникационных систем* («*РАДИОИНФОКОМ-2022*»): *Сб. научных статей по материалам VI Международной научно-практической конференции*. М.: РТУ МИРЭА; 2022. С. 72–75. https://www.elibrary.ru/knywzk

About the Authors

Alexander V. Ksendzuk, Dr. Sci. (Eng.), Head of the Department of Radioelectronic Systems, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: ks_alex@mail.ru. Scopus Author ID 56628472300, RSCI SPIN-code 2389-6036, https://orcid.org/0009-0001-7084-1433, https://www.researchgate.net/profile/Alexander-Ksendzuk-2

Ivan A. Kuznetsov, Postgraduate Student, Public Joint Stock Company VYMPEL Interstate Corporation (10-1, Geroyev Panfilovtsev ul., Moscow, 125480 Russia). E-mail: 0601ivankuznetsov@gmail.com. https://orcid.org/0009-0009-0045-6626

Об авторах

Ксендзук Александр Владимирович, д.т.н., заведующий кафедрой № 346 – радиоэлектронных систем, Институт радиоэлектроники и информатики, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: ks_alex@mail.ru. Scopus Author ID 56628472300, SPIN-код РИНЦ 2389-6036, https://orcid.org/0009-0001-7084-1433, https://www.researchgate.net/profile/Alexander-Ksendzuk-2

Кузнецов Иван Алексеевич, аспирант, ПАО «МАК «Вымпел» (125480, Россия, Москва, ул. Героев Панфиловцев, д. 10, к. 1). E-mail: 0601ivankuznetsov@gmail.com. https://orcid.org/0009-0009-0045-6626

Translated from Russian into English by L. Bychkova
Edited for English language and spelling by Thomas A. Beavitt

Mathematical modeling

Математическое моделирование

UDC 537.811
https://doi.org/10.32362/2500-316X-2025-13-5-87-94
EDN SPEPVZ

RESEARCH ARTICLE

Analytical model for the normal component of magnetic induction of a permanent magnets

Mikhail M. Zakatov @

Dmitriy Mikhailik Academy of Civil Defence of the Ministry of Emergency Situations of Russia, Khimki, 141435 Russia
@ *Corresponding author, e-mail: zakatov46@mail.ru*

• **Submitted:** 10.02.2025 • **Revised:** 09.04.2025 • **Accepted:** 21.07.2025

Abstract

Objectives. In a measuring system based on the inductive transmission of information from a moving structure to a stationary signal receiver, the signal carrying useful information about the parameters of the moving structure is formed by a magnetic system containing a permanent magnet mounted on the stationary part of the measuring system. The magnetic field of the permanent magnet (MFPM) determines the magnetic flux, and, consequently, the induction current in a conducting coil located on the moving structure. In order to theoretically justify the parameters of the measuring system including the optimization of its components, a simple and easy-to-use analytical model of the useful signal for determining the requirements for the mathematical description of the MFPM is required. The use of known solutions for developing an analytical model of the useful signal of the measuring system is complicated by the need to use inverse trigonometric functions or the results of numerical calculations. The present work sets out to obtain an exact solution to the problem of calculating the MFPM and on this basis to develop a simple, convenient analytical model of the normal component of the magnetic induction vector (NCMIV) of a permanent magnet used to develop an analytical model of the useful signal.

Methods. The equivalent solenoid method was used along with mathematical analysis approaches.

Results. An exact solution for calculating the normal component of the magnetic induction vector of the parallelepiped-shaped permanent magnet was obtained. Based on this, a straightforward and easy-to-use analytical model of the NCMIV was developed, which closely approximates the formula derived for the exact solution.

Conclusions. The developed analytical model of the NCMIV can be used for theoretical development of an analytical model of the useful signal of a measuring system with inductive transmission of information about the parameters of a moving structure to a stationary signal receiver.

Keywords: permanent magnet, magnetic induction, equivalent solenoid, normal component, analytical model

For citation: Zakatov M.M. Analytical model for the normal component of magnetic induction of a permanent magnet. *Russian Technological Journal.* 2025;13(5):87–94. https://doi.org/10.32362/2500-316X-2025-13-5-87-94, https://www.elibrary.ru/SPEPVZ

Financial disclosure: The author has no financial or proprietary interest in any material or method mentioned.

The author declares no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Аналитическая модель нормальной составляющей магнитной индукции постоянного магнита

М.М. Закатов [@]

Академия гражданской защиты МЧС России имени Д.И. Михайлика, Химки, 141435 Россия
[@] *Автор для переписки, e-mail: zakatov46@mail.ru*

• **Поступила:** 10.02.2025 • **Доработана:** 09.04.2025 • **Принята к опубликованию:** 21.07.2025

Резюме

Цели. В измерительной системе с индукционной передачей информации с перемещающейся конструкции на неподвижный приемник информационный сигнал, несущий информацию о параметрах перемещающейся конструкции, формируется магнитной системой, содержащей постоянный магнит, установленный на неподвижной части измерительной системы. Магнитное поле постоянного магнита (МППМ) определяет магнитный поток, и, следовательно, индукционный ток в другом элементе магнитной системы – проводящем витке, расположенном на перемещающейся конструкции. Для теоретического обоснования параметров измерительной системы, в т.ч. для оптимизации ее составных частей, необходима простая, удобная для применения аналитическая модель информационного сигнала (АМИС), что определяет требования к математическому описанию МППМ. Известные решения задач по расчету МППМ содержат обратные тригонометрические функции или представлены результатами численных расчетов, что затрудняет их использование для разработки АМИС измерительной системы. Целью данной статьи является получение точного решения задачи расчета МППМ и разработка на основании этого точного решения аналитической модели нормальной составляющей вектора магнитной индукции (НСВМИ) постоянного магнита, используемой для разработки АМИС.

Методы. Использовались методы математического анализа и метод эквивалентного соленоида.

Результаты. Получено точное решение задачи расчета НСВМИ МППМ, имеющего форму параллелепипеда, на основании которого получено выражение, аппроксимирующее формулу точного решения, – аналитическая модель НСВМИ.

Выводы. Полученная аналитическая модель НСВМИ может быть использована для теоретической разработки АМИС измерительной системы с индукционной передачей информации о параметрах перемещающейся конструкции на неподвижный приемник сигнала.

Ключевые слова: постоянный магнит, магнитная индукция, эквивалентный соленоид, нормальная составляющая, аналитическая модель

Для цитирования: Закатов М.М. Аналитическая модель нормальной составляющей магнитной индукции постоянного магнита. *Russian Technological Journal*. 2025;13(5):87–94. https://doi.org/10.32362/2500-316X-2025-13-5-87-94, https://www.elibrary.ru/SPEPVZ

Прозрачность финансовой деятельности: Автор не имеет финансовой заинтересованности в представленных материалах или методах.

Автор заявляет об отсутствии конфликта интересов.

INTRODUCTION

A radio channel is mostly used to convey information about the parameters of a structure that changes position relative to a stationary signal receiver (consider, for instance, relaying gas pressure readings within a vehicle's tire). In this situation, measurement transducers are mounted on the structure to transform characteristics defining its status into electrical impulses, which are subsequently transmitted to a radio transmitter. A radio signal receiver is mounted on the stationary part, such as the body of the vehicle, close to where the pneumatic tire wheel changes position.

Transmitting information via a radio channel often encounters challenges stemming from limitations imposed on the use of electromagnetic radiation. In

a measuring system with inductive transmission of information about the state of a moving structure [1, 2], the useful signal carrying information about the state of the structure is formed by a magnetic system, which is formalized as shown in Fig. 1. A rectangular conducting coil (1) is positioned on the upper edge of a permanent magnet (PM), which is configured as a parallelepiped with a height of H and the base having dimensions a and b. In a plane parallel to the PM upper edge, a structure with a placed-in rectangular coil 2 travels relative to the magnet at velocity V, with the center of coil 2 (point O_2) projected onto the Ox axis in a manner that its projection (point A) traverses along the Ox axis. The distance between the plane and the upper edge of the magnet is equal to d. The electrical resistance of coil 2 is dictated by the design specifications.

When coil 2 travels within the magnetic field of a permanent magnet (MFPM), an induced electric current arises within coil 2. The intensity of this current is influenced by the velocity of coil 2, as well as its internal electrical resistance and the normal component of the magnetic induction vector (NCMIV) of MFPM perpendicular to the plane of coil 2. This induced current in moving coil 2 subsequently produces an alternating magnetic field around coil 1. The electromotive force of coil 1 resulting from the magnetic flux of the alternating magnetic field created by the induced current in moving coil 2 acts as a signal for carrying useful information about the system parameters.

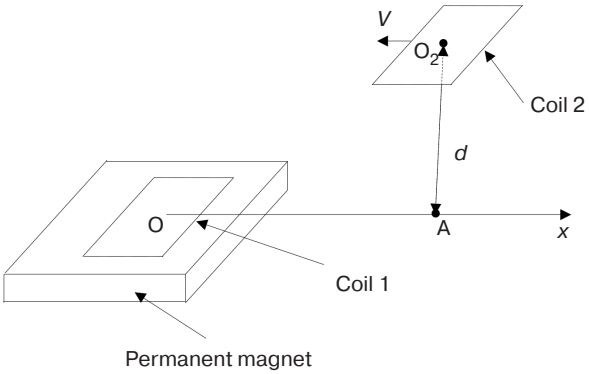


Fig. 1. Schematic of the magnetic system

Various analytical [3–11] and numerical [12–24] techniques have been established for analyzing and calculating the properties of magnetic fields within diverse magnetic systems. The present study employs the equivalent solenoid method to precisely calculate the NCMIV distribution within the MFPM in the plane parallel to the PM upper edge. These findings are then used to create the NCMIV analytical model, which can be used to develop analytical models for magnetic fluxes responsible for generating induced electric currents in conducting coils 1 and 2. The findings regarding the MFPM distribution, as detailed in [11], are expanded upon by additional calculations focusing on the MFPM distribution in the geometric area near permanent magnet.

PROBLEM STATEMENT

Figure 2 depicts the magnetic field produced by a permanent magnet that is uniformly magnetized along the Oz axis. The origin of the coordinate system $Oxyz$ is placed at the magnet base center. Based on Fig. 2, we calculate the normalized MFPM NCMIV, $B_{z\,n}(x, y, z)$, at point M(x, y, z) located in the specified plane $z \geq H$.

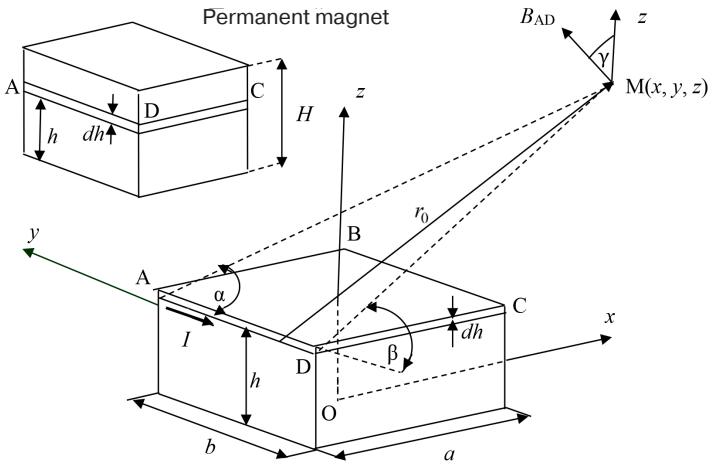


Fig. 2. Geometric layout for calculating MFPM NCMIV

Utilizing the equivalent solenoid approach [3–6], we substitute the permanent magnet with a system of surface electric currents, characterized by a linear density λ, flowing in planes perpendicular to the Oz axis. Then, a surface electric current of strength $I = \lambda dh$ flows along a strip spanning the PM side surface, with a width of dh and situated at

a height h above its base. The strip perimeter is defined by the segments AD, DC, CB, and BA, where |AD| = |CB| = b, |DC| = |BA| = a.

The magnetic field strength, $B_{AD}(x, y, z)$, created by current I flowing through segment AD at point M(x, y, z) is determined by the following equation [3]:

$$B_{AD}(x, y, z) = \frac{\mu_0}{4\pi} I \frac{\cos\alpha - \cos\beta}{r_0}, \tag{1}$$

where r_0 is the distance from point M(x, y, z) to segment AD.

The magnetic induction vector B_{AD} is directed at an angle γ to the Oz axis.

Based on Fig. 2, the following relationships are derived:

$$-\cos\beta = \frac{y + \frac{b}{2}}{\sqrt{\left(x + \frac{a}{2}\right)^2 + \left(y + \frac{b}{2}\right)^2 + (z - h)^2}}, \tag{2}$$

$$\cos\alpha = \frac{\frac{b}{2} - y}{\sqrt{\left(x + \frac{a}{2}\right)^2 + \left(y - \frac{b}{2}\right)^2 + (z - h)^2}}, \tag{3}$$

$$\cos\gamma = \frac{x + \frac{a}{2}}{\sqrt{\left(x + \frac{a}{2}\right)^2 + (z - h)^2}}. \tag{4}$$

The NCMIV B_{ADz} generated by the electric current flowing through segment AD is determined by the following equation:

$$B_{ADz}(x, y, z) = B_{AD}(x, y, z)\cos\gamma. \tag{5}$$

The NCMIV arising from the electric currents within segments DC, CB, and BA are determined in the same way.

The NCMIV produced by the electric current I flowing through a strip of width dh at a height h above PM is the sum of NCMIVs generated by the currents within segments AD, DC, CB, and BA:

$$B_{hz}(x, y, z) = B_{ADz}(x, y, z) + B_{DCz}(x, y, z) + B_{CBz}(x, y, z) + B_{BAz}(x, y, z). \tag{6}$$

In the following, the AD, DC, CB, and BA symbols denote the PB side surfaces containing corresponding segments.

Taking into account Eqs. (1)–(6), the formula for calculating the normalized NCMIV of the electric current on the side surface AD is derived as follows:

$$\begin{aligned}\frac{4\pi}{\mu_0\lambda}\int_0^H B_{ADz}(x, y, z)dh = &\left(\arcsin\frac{\alpha_{11}z}{\sqrt{(\alpha_{11}^2 + \beta_1^2)(\beta_1^2 + z^2)}} - \arcsin\frac{\alpha_{11}(z - H)}{\sqrt{(\alpha_{11}^2 + \beta_1^2)(\beta_1^2 + (z - H)^2)}}\right) + \\ &+ \left(\arcsin\frac{\alpha_{12}z}{\sqrt{(\alpha_{12}^2 + \beta_1^2)(\beta_1^2 + z^2)}} - \arcsin\frac{\alpha_{12}(z - H)}{\sqrt{(\alpha_{12}^2 + \beta_1^2)(\beta_1^2 + (z - H)^2)}}\right),\end{aligned} \tag{7}$$

where $\alpha_{11} = \frac{b}{2} - y$, $\alpha_{12} = \frac{b}{2} + y$, $\beta_1 = x + \frac{a}{2}$.

The normalized MFPM NCMIV, $B_{z\,n}(x, y, z)$, equals the sum of the normalized NCMIVs of all electric currents on the PM side surfaces. Taking into account (7), we obtain the following:

$$B_{z\,\mathrm{n}}(x,\,y,\,z)=\sum_{i=1}^{4}\left(\left(\arcsin\frac{\alpha_{i1}z}{\sqrt{(\alpha_{i1}^2+\beta_i^2)(\beta_i^2+z^2)}}-\arcsin\frac{\alpha_{i1}(z-H)}{\sqrt{(\alpha_{i1}^2+\beta_i^2)(\beta_i^2+(z-H)^2)}}\right)+\right.$$
$$\left.+\left(\arcsin\frac{\alpha_{i2}z}{\sqrt{(\alpha_{i2}^2+\beta_i^2)(\beta_i^2+z^2)}}-\arcsin\frac{\alpha_{i2}(z-H)}{\sqrt{(\alpha_{i2}^2+\beta_i^2)(\beta_i^2+(z-H)^2)}}\right)\mathrm{sign}(\beta_i)\right), \quad (8)$$

where $\alpha_{11}=\frac{b}{2}-y$, $\alpha_{12}=\frac{b}{2}+y$, $\beta_1=x+\frac{a}{2}$ (for AD); $\alpha_{21}=x+\frac{a}{2}$, $\alpha_{22}=\frac{a}{2}-x$, $\beta_2=y+\frac{b}{2}$ (for DC); $\alpha_{31}=y+\frac{b}{2}$, $\alpha_{32}=\frac{b}{2}-y$, $\beta_3=\frac{a}{2}-x$ (for CB); $\alpha_{41}=\frac{a}{2}-x$, $\alpha_{42}=\frac{a}{2}+x$, $\beta_4=\frac{b}{2}-y$ (for BA).

The NCMIV of the electric current on the CB side surface changes sign at $x>\frac{a}{2}$, similar to the NCMIV of the electric current on the BA side surface at $y>\frac{b}{2}$. The change of NCMIV signs is taken into account by the function.

Since the function $B_{z\,\mathrm{n}}(x, y, z)$ is even with respect to both x and y coordinates, determining its distribution within a specified plane only requires evaluating it in one quarter as limited by the inequalities $0\le x$ and $0\le y$. To simplify $B_{z\,\mathrm{n}}(x, y, z)$ for further investigations, formula (8) is decomposed into a Taylor series. The analysis indicates that formula (8) can be replaced by the more concise and computationally efficient NCMIV analytical model as follows:

$$B_{z\,\mathrm{n.anal}}(x,y,z)=\begin{cases}B_{z\,\mathrm{n}}(0,\,0,\,z)\left[\left(\frac{y}{y_{\max}}\right)^2-1\right]^2\left[\left(\frac{x}{x_{\max}}\right)^2-1\right]^2 & \text{at } |x|\le x_{\max},\ |y|\le y_{\max},\\ 0 & \text{at } |x|>x_{\max},\ |y|>y_{\max},\end{cases} \quad (9)$$

where $x_{\max}$, $y_{\max}$ are the "dimensions" of the magnetic field in the plane z = const, i.e., the coordinates $x_{\max}$, $y_{\max}$, which depend on the coordinate z, are chosen so that the MFPM NCMIV becomes equal to zero in the region $|x|\ge x_{\max}$, $|y|\ge y_{\max}$ of the plane z = const.

In Fig. 3, the graphs of the dependencies of the normalized MFPM NCMIV, $B_{z\,\mathrm{n}}(x, y, z)$, in the plane z = 13 mm are presented as a function of the x-coordinate at various y-coordinate values, with the following PM parameters: $a=b$ = 16 mm, H = 8 mm. The values of $B_{z\,\mathrm{n}}(x, y, z)$ are calculated using the exact formula (8) for graphs *1*, *2*, and *3* and formula (9) for graphs *4*, *5*, and *6*.

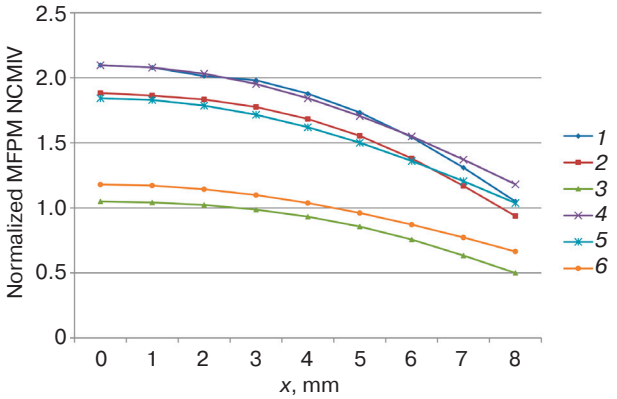


Fig. 3. Graphs of the normalized MFPM NCMIV: *1* is $B_{z\,\mathrm{n}}(x, 0, 13)$, *2* is $B_{z\,\mathrm{n}}(x, 4, 13)$, *3* is $B_{z\,\mathrm{n}}(x, 8, 13)$, *4* is $B_{z\,\mathrm{n.anal}}(x, 0, 13)$, *5* is $B_{z\,\mathrm{n.anal}}(x, 4, 13)$, and *6* is $B_{z\,\mathrm{n.anal}}(x, 8, 13)$

The analysis reveals a significant difference between the calculation results using the exact formula (8) and those using the approximation formula (9) for points in the plane with coordinates $x\ge 6$ mm, $y\ge 6$ mm. The relative error, δ, of the MFPM NCMIV calculations based on the analytical model belongs to the segment $15\%\le\delta\le 33\%$.

For points on the plane with coordinates $x\le 6$ mm, $y\le 6$ mm, the relative error δ is less than 15%. Moreover, for points on the plane limited by the inequality $r\le 4$ mm, where r is the distance between the coordinate system origin and the point on the plane, the relative error is $\delta\le 3.8\%$; for points of the plane limited by the inequality $r\le 2$ mm, the relative error is $\delta\le 2.6\%$.

The parameters of the useful signal of the measuring system, including its amplitude and time-dependent shape over its duration, are determined by the rate of change of magnetic fluxes within coils 1 and 2 when the center of coil 2 is above the center of PM cross-section. Since the magnetic fluxes and the rate of their changes are primarily determined by the central part of the MFPM NCMIV as represented by the points in the plane with coordinates limited by the inequalities $x \leq 6$ mm and $y \leq 6$ mm, the error in calculating the useful signal parameters in this case is minimal.

At high values of the MFPM NCMIV coordinates, the MFPM peripheral part decreases rapidly, while the rate of change of the magnetic fluxes determined by the MFPM peripheral part is less than that determined by the MFPM central part.

Error estimations derived from NCMIV calculations employing the approximation formula (9) reveal the potential of the analytical NCMIV model for creating an analytical model for the useful signal of the measuring system. This analytical model achieves a calculation error for useful signal values not exceeding 15%.

CONCLUSIONS

Based on the results of the research, the following can be deduced:

1. An analytical formula for calculating the NCMIV of the parallelepiped-shaped MFPM, which is oriented perpendicular to a specified plane aligned with the PM end face, has been derived by applying the equivalent solenoid technique.
2. The obtained expression for the NCMIV analytical model enables straightforward calculations of the magnetic fluxes within coils 1 and 2.
3. The feasibility of applying the NCMIV analytical model to theoretically develop an analytical model for the useful signal of the measuring system with inductive signal transmission is confirmed.

REFERENCES

1. Zakatov M.M. Theoretical foundations of an information system with inductive signal transmission. *Radiopromyshlennost' = Radio Industry*. 2001;1:29–39 (in Russ.).
2. Zakatov M.M. *Passive Automotive Tire Air Pressure and Temperature Meter*: RF Pat. 2509656. Publ. 20.03.2014 (in Russ.).
3. Sivukhin D.V. *Obshchii kurs fiziki: Elektrichestvo* (*General Course of Physics: in 5 v. V. 3. Electricity*). Moscow: Fizmatlit; 2004. 656 p. (in Russ.).
4. Kustov M.S., Druina D.V., Mikhailova O.O., Polyakov I.G. Magnetic field calculation for uniformly magnetized bodies by the methods of equivalent solenoid and magnetic charges. *Vestnik TvGU. Seriya: Fizika = Bulletin of TvGU. Series: Physics*. 2010;8:17–35 (in Russ.).
5. Erlichson H. The magnetic field of a circular turn. *Am. J. Phys.* 1989;57(7):607–608. https://doi.org/10.1119/1.15955
6. Vorontsov A.A., Slesarev Yu.N., Karpushin E.V. Mathematical modeling of magnetic fields of two-axis magnetostrictive inclinometers containing constant magnet shaped as rectangular parallelepiped. *Vestnik TGTU = Transactions of the TSTU.* 2013;19(1):25–29 (in Russ.).
7. Vorontsov A.A., Slesarev Yu.N., Karpushin E.V. Mathematical modeling of magnetic fields of two-coordinate magnetostrictive tiltmeters containing a ring or solid permanent magnet. *Izvestiya Penzenskogo gosudarstvennogo pedagogicheskogo universiteta im. V.G. Belinskogo.* 2012;30:467–472 (in Russ.).
8. Ravaud R., Lemarquand G., Depollier C., et al. Analytical calculation of the magnetic field created by permanent-magnet rings. *IEEE Transact. Magnet*. 2008;44(8):1982–1989. https://doi.org/10.1109/TMAG.2008.923096
9. Vagin D.V., Gerasimenko T.N., Polyakov P.A. Exact analytical expression for magnetic field induction of rectangular shape sample. *Moscow Univ. Phys.* 2008;63(6):422–422. https://doi.org/10.3103/S0027134908060131
[Original Russian Text: Vagin D.V., Gerasimenko T.N., Polyakov P.A. Exact analytical expression for magnetic field induction of rectangular shape sample. *Vestnik Moskovskogo universiteta. Seriya 3: Fizika. Astronomiya.* 2008;6:53–55 (in Russ.).]
10. Xiao-Fan G., Yong Y., Xiao-Jing Z. Analytic expression of magnetic field distribution of rectangular permanent magnets. *Appl. Math. Mech.* 2004;25:297–306. https://doi.org/10.1007/BF02437333
11. Artamonov A.N., Erzhenin N.R., Zakatov M.M. Analytical model of the normal component of the magnetic induction vector. In: *Proceedings of the 27th International Scientific-Practical Conference* "*Prevention. Salvation. Help.*" Khimki; 2017. P. 5–9 (in Russ.).
12. Rezinkina M.M. Parameters of thin electromagnetic shields that provide a decrease in magnetic induction. *Tech. Phys*. 2014;59(2):155–161. https://doi.org/10.1134/S1063784214020194
[Original Russian Text: Rezinkina M.M. Parameters of thin electromagnetic shields that provide a decrease in magnetic induction. *Zhurnal tekhnicheskoi fiziki*. 2014;84(2):1–7 (in Russ.).]

13. Rezinkina M.M. Numerical calculation of the magnetic field and magnetic moment of ferromagnetic bodies of complex spatial configuration. *Tech. Phys*. 2009;54(8):1092–1101. https://doi.org/10.1134/S1063784209080027
[Original Russian Text: Rezinkina M.M. Numerical calculation of the magnetic field and magnetic moment of ferromagnetic bodies of complex spatial configuration. *Zhurnal tekhnicheskoi fiziki*. 2009;79(8):8–17 (in Russ.).]
14. Slesarev Yu.N., Martens-Atyusheva K.Yu., Rodionov S.V. Research of mathematical models of magnetic fields of two-coordinate magnetostriction tiltmeters. *Modeli, sistemy, seti v ekonomike, tekhnike, prirode i obshchestve = Models, Systems, Networks in Economics, Technology, Nature and Society.* 2016;2(18):290–298 (in Russ.).
15. Bulyzhev E.M., Menshov E.N., Dzhavakhiya G.A. Modeling the field permanent magnet. *Izvestiya Samarskogo nauchnogo tsentra RAN = Izvestiya of Samara Scientific Center of the Russian Academy of Sciences*. 2011;13(4):106–110 (in Russ.).
16. Cherkasova O.A. Research of the magnetic field of the permanent magnet by means of computer modeling. *Geteromagnitnaya mikroelektronika = Heteromagnetic Microelectronics.* 2014;17:112–120 (in Russ.).
17. Matyuk V.F., Osipov A.A., Strelyukhin A.V. Magnetization distribution along the cylindrical rod in longitudinal constant uniform magnetic field. *Russ. Electr. Eng*. 2009;80(8):450–458. https://doi.org/10.3103/S1068371209080082
[Original Russian Text: Matyuk V.F., Osipov A.A., Strelyukhin A.V. Magnetization distribution along the cylindrical rod in longitudinal constant uniform magnetic field. *Elektrotekhnika.* 2009;8:37–46 (in Russ.).]
18. Andreeva E.G., Tatevosyan A.A., Semina I.A. Research of the axisymmetric model of an open-type magnetic system. *Omskii nauchnyi vestnik = Omsk Scientific Bulletin.* 2010;1(87):110–113 (in Russ.).
19. Lagutin A.S., Grigoriev G.Yu. Magnetic system of the xenon core polarizer. *Prikladnaya fizika = Appl. Phys*. 2023;3:92–95 (in Russ.).
20. Alievskii B.L., Orlov V.L. *Raschet parametrov magnitnykh polei osesimmetrichnykh katushek* (*Calculation of Magnetic Field Parameters of Axisymmetric Coils*). Moscow: Energoatomizdat; 1983. 112 p. (in Russ.).
21. Camacho J.M., Sosa V. Alternative method to calculate the magnetic field of permanent magnets with azimuthal symmetry. *Revista Mexicana de Física E*. 2013;59(1):8–17. https://doi.org/10.31349/RevMexFisE.21.020701
22. Semykina I.Yu., Zav'yalov V.M., Dubkov E.A., Uglova M.B. Analysis of the magnetic field induction distribution in the vicinity of a rectangular coil profile of a wireless power transmission system. In: *Proceedings of the International Scientific and Practical Conference* "*Advanced Technologies and Materials.*" Sevastopol. 2022. P. 235– 239 (in Russ.).
23. Li B., Zhang J., Zhao X., Liu B., Dong H. Research on air gap magnetic field characteristics of trapezoidal Halbach permanent magnet linear synchronous motor based on improved equivalent surface current. *Energies*. 2023;16(2):793–805. https://doi.org/10.3390/en16020793
24. Osipov V.V., Orlov A.N., Lisenkov V.V., et al. Distribution of the Magnetic Field in the Gap between Two Permanent Magnets: Calculated and Experimental Data and Their Applications. *Instrum. Exp. Tech.* 2023;66(6):995–1002. https://doi.org/10.1134/S0020441223060064
[Original Russian Text: Osipov V.V., Orlov A.N., Lisenkov V.V., Maksimov R.N., Shitov V.A. Distribution of the Magnetic Field in the Gap between Two Permanent Magnets: Calculated and Experimental Data and Their Applications. *Pribory i tekhnika eksperimenta*. 2023;6:111–118 (in Russ.). https://doi.org/10.31857/S003281622306006X]

СПИСОК ЛИТЕРАТУРЫ

1. Закатов М.М. Теоретические основы информационной системы с индукционной передачей сигнала. *Радиопромышленность.* 2001;1:29–39.
2. Закатов М.М. *Пассивный измеритель давления и температуры воздуха в шине колеса автомобиля*: пат. 2509656 РФ. Заявка № 2012140219/11; заявл. 20.09.2012, опубл. 20.03.2014. Бюл. 8.
3. Сивухин Д.В. *Общий курс физики: в 5 т. Т. 3. Электричество*. М.: Физматлит; 2004. 656 с.
4. Кустов М.С., Друина Д.В., Михайлова О.О., Поляков И.Г. Расчет магнитных полей однородно намагниченных тел методами эквивалентного соленоида и магнитных зарядов. *Вестник ТвГУ. Серия «Физика».* 2010;8:17–35.
5. Erlichson H. The magnetic field of a circular turn. *Am. J. Phys.* 1989;57(7):607–608. https://doi.org/10.1119/1.15955
6. Воронцов А.А., Слесарев Ю.Н., Карпухин Э.В. Математическое моделирование магнитных полей двухкоординатных магнитострикционных наклономеров, содержащих постоянный магнит в форме прямоугольного параллелепипеда. *Вестник Тамбовского ГТУ*. 2013;19(1):25–29.
7. Воронцов А.А., Слесарев Ю.Н., Карпухин Э.В. Математическое моделирование магнитных полей двухкоординатных магнитострикционных наклономеров, содержащих кольцевой или сплошной постоянный магнит. *Известия Пензенского государственного педагогического университета им. В.Г. Белинского.* 2012;30:467–472.
8. Ravaud R., Lemarquand G., Depollier C., et al. Analytical calculation of the magnetic field created by permanent-magnet rings. *IEEE Transact. Magnet*. 2008;44(8):1982–1989. https://doi.org/10.1109/TMAG.2008.923096
9. Вагин Д.В., Герасименко Т.Н., Поляков П.А. Точное аналитическое выражение для индукции магнитного поля образца прямоугольной формы. *Вестник Московского университета. Серия 3: Физика. Астрономия.* 2008;6:53–55.
10. Xiao-Fan G., Yong Y., Xiao-Jing Z. Analytic expression of magnetic field distribution of rectangular permanent magnets. *Appl. Math. Mech.* 2004;25:297–306. https://doi.org/10.1007/BF02437333
11. Артамонов А. Н., Ерженин Н.Р., Закатов М.М. Аналитическая модель нормальной составляющей вектора магнитной индукции. В сб.: *Материалы XXVII Международной научно-практической конференции «Предупреждение, спасение, помощь».* Химки; 2017. С. 5–9.

12. Резинкина М.М. Выбор параметров тонких электромагнитных экранов для снижения уровней магнитной индукции. *Журн. техн. физики*. 2014;84(2):1–7.
13. Резинкина М.М. Численный расчет магнитного поля и магнитного момента ферромагнитных тел сложной пространственной конфигурации. *Журн. техн. физики.* 2009;79(8):8–17.
14. Слесарев Ю.Н., Мартенс-Атюшева К.Ю., Родионов С.В. Исследование математических моделей магнитных полей двухкоординатных магнитострикционных наклономеров. *Модели, системы, сети в экономике, технике, природе и обществе.* 2016;2(18):290–298.
15. Булыжев Е.М., Меньшов Е.Н., Джавахия Г.А. Моделирование поля постоянного магнита *Известия Самарского научного центра РАН*. 2011;13(4):106–110.
16. Черкасова О.А. Исследование магнитного поля постоянного магнита с помощью компьютерного моделирования. *Гетеромагнитная микроэлектроника.* 2014;17:112–120.
17. Матюк В.Ф., Осипов А.А., Стрелюхин А.В. Распределение намагниченности вдоль цилиндрического стержня, находящегося в продольном постоянном однородном магнитном поле. *Электротехника.* 2009;8:37–46.
18. Андреева Е.Г., Татевосян А.А., Семина И.А. Исследование осесимметричной модели магнитной системы открытого типа. *Омский научный вестни*к. 2010;1(87):110–113.
19. Лагутин А.С., Григорьев Г.Ю. Магнитная система поляризатора ядер ксенона. *Прикладная физика.* 2023;3:92–95.
20. Алиевский Б.Л., Орлов В.Л. *Расчет параметров магнитных полей осесимметричных катушек.* М.: Энергоатомиздат; 1983. 112 с.
21. Camacho J.M., Sosa V. Alternative method to calculate the magnetic field of permanent magnets with azimuthal symmetry. *Revista Mexicana de Física E*. 2013;59(1):8–17. https://doi.org/10.31349/RevMexFisE.21.020701
22. Семыкина И.Ю., Завьялов В.М., Дубков Е.А., Углова М.Б. Анализ распределения индукции магнитного поля в окрестностях катушки прямоугольного профиля системы беспроводной передачи энергии. В сб.: *Перспективные технологии и материалы: Материалы Международной научно–практической конференции.* Севастополь. 2022. С. 235–239.
23. Li B., Zhang J., Zhao X., Liu B., Dong H. Research on air gap magnetic field characteristics of trapezoidal Halbach permanent magnet linear synchronous motor based on improved equivalent surface current method. *Energies*. 2023;16(2):793–805. https://doi.org/10.3390/en16020793
24. Осипов В.В., Орлов А.Н., Лисенков В.В., Максимов Р.Н., Шитов В.А. Распределение магнитного поля в зазоре между двумя постоянными магнитами: расчетные и экспериментальные данные и их приложения. *Приборы и техника эксперимента*. 2023;6:111–118. https://doi.org/10.31857/S003281622306006X

About the Author

Mikhail M. Zakatov, Cand. Sci. (Eng.), Senior Researcher, Associate Professor, Department of Mechanics and Engineering Graphics, Civil Defence Academy EMERCOM of Russia (1, Sokolovskaya ul., mkr. Novogorsk, Khimki, Moscow oblast, 141435 Russia). E-mail: zakatov46@mail.ru. https://orcid.org/0009-0006-1249-8039

Об авторе

Закатов Михаил Михайлович, к.т.н., старший научный сотрудник, доцент, кафедра механики и инженерной графики, ФГБВОУ ВО «Академия гражданской защиты Министерства Российской Федерации по делам гражданской обороны, чрезвычайным ситуациям и ликвидации последствий стихийных бедствий имени генерал-лейтенанта Д.И. Михайлика» (ФГБВОУ ВО «Академия гражданской защиты МЧС России») (141435, Россия, Московская обл., г.о. Химки, мкр. Новогорск, ул. Соколовская, стр. 1). E-mail: zakatov46@mail.ru. https://orcid.org/0009-0006-1249-8039

Translated from Russian into English by K. Nazarov
Edited for English language and spelling by Thomas A. Beavitt

Mathematical modeling

Математическое моделирование

UDC 538.93
https://doi.org/10.32362/2500-316X-2025-13-5-95-103
EDN YKAHQQ

RESEARCH ARTICLE

Heat transfer in a porous medium having an ordered gyroid-based macrostructure

Andrey I. Popov [@], **Anton V. Eremin**

Samara State Technical University, Samara, 443100 Russia
[@] *Corresponding author, e-mail: popov.ai@samgtu.ru*

• **Submitted:** 27.12.2024 • **Revised:** 10.03.2025 • **Accepted:** 23.07.2025

Abstract

Objectives. Triply periodic minimal surfaces are non-intersecting surfaces with zero mean curvature, consisting of elements repeating in three directions of the Cartesian coordinate system. The use of structures based on minimal surfaces in heat engineering equipment is associated with their advantages over classical lattice and honeycomb structures, often used in practice. The aim of the work is to study heat transfer during filtration flow in a porous medium of an incompressible fluid having an ordered macrostructure based on gyroid triply periodic minimal surface.

Methods. In order to solve the problem of heat transfer in a porous medium, the finite difference method is used. As a means of implementing the finite difference method algorithm, the *Heat Transfer Solver* software was developed in the Python programming language.

Results. The described software program was used to obtain a numerical solution of the heat transfer problem in a porous medium with an ordered macrostructure using the finite difference method. The program functionality enables the investigation of the heat transfer process dynamics and the influence of various parameters on the temperature distribution. The program was used to study the heat transfer process in a porous medium based on gyroid triply periodic minimal surface. Graphical dependencies of the solid framework and fluid temperatures, as well as the heat flux on the coordinate at different time steps, were obtained. Characteristic time intervals with the highest absolute temperature gradient values were identified.

Conclusions. The results of the work, including both the developed software and the obtained temperature dependencies, can be used in a number of engineering problems where it is important to predict the temperature distribution in porous materials under various operating conditions.

Keywords: porous medium, fluid flow, heat transfer, triply periodic minimal surface, gyroid, finite difference method

For citation: Popov A.I., Eremin A.V. Heat transfer in a porous medium having an ordered gyroid-based macrostructure. *Russian Technological Journal.* 2025;13(5):95–103. https://doi.org/10.32362/2500-316X-2025-13-5-95-103, https://www.elibrary.ru/YKAHQQ

Financial disclosure: The authors have no financial or proprietary interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Исследование теплопереноса в пористой среде с упорядоченной макроструктурой на основе гироида

А.И. Попов [@], **А.В. Еремин**

Самарский государственный технический университет, Самара, 443100 Россия
[@] *Автор для переписки, e-mail: popov.ai@samgtu.ru*

• **Поступила:** 27.12.2024 • **Доработана:** 10.03.2025 • **Принята к опубликованию:** 23.07.2025

Резюме

Цели. Трижды периодические минимальные поверхности – это непересекающиеся поверхности с нулевой средней кривизной, состоящие из повторяющихся в трех направлениях декартовой системы координат элементов. Использование конструкций, основанных на минимальных поверхностях, в теплотехническом оборудовании связано с их преимуществами перед классическими решетчатыми и сотовыми конструкциями, часто применяемыми на практике. Целью работы является исследование теплопереноса при фильтрационном течении несжимаемой жидкости в пористой среде с упорядоченной макроструктурой на основе трижды периодической минимальной поверхности (гироида).

Методы. Для решения задачи теплопереноса в пористой среде применяется метод конечных разностей. Для реализации алгоритма метода конечных разностей разработано программное обеспечение *Heat Transfer Solver* на языке программирования Python.

Результаты. В рамках настоящего исследования разработано программное обеспечение на языке программирования Python для численного решения методом конечных разностей задачи теплопереноса в пористой среде с упорядоченной макроструктурой. Функционал программы позволяет исследовать динамику процесса теплопереноса и влияние различных параметров на распределение температуры. При помощи данной программы изучен процесс теплопереноса в пористой среде на основе гироида. Получены графические зависимости температуры твердотельного каркаса и жидкости, а также теплового потока от координаты в различные моменты времени. Определены характерные временные интервалы, в которых наблюдается наибольшее абсолютное значение градиента температур.

Выводы. Результаты работы, включающие как разработанное программное обеспечение, так и зависимости температур, могут найти применение в ряде инженерных задач, где важным является прогнозирование температурного распределения в пористых материалах при различных условиях эксплуатации.

Ключевые слова: пористая среда, течение жидкости, теплообмен, трижды периодическая минимальная поверхность, гироид, метод конечных разностей

Для цитирования: Попов А.И., Еремин А.В. Исследование теплопереноса в пористой среде с упорядоченной макроструктурой на основе гироида. *Russian Technological Journal*. 2025;13(5):95–103. https://doi.org/10.32362/2500-316X-2025-13-5-95-103, https://www.elibrary.ru/YKAHQQ

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

At present, an important task in theoretical and applied heat engineering is associated with the need to increase the energy efficiency of various heat engineering equipment [1] used in a wide range of industries. One of the methods for intensifying heat exchange consists in the use of porous materials in the design of heat exchange devices [2], catalysts [3], etc. Such porous materials may be used for increasing the heat exchange surface area and flow turbulization in the heat exchanger paths.

Most porous materials used in practice have a stochastic structure, i.e., the pores have a random shape and size. However, there are whole classes of porous materials having an ordered structure, such as lattice and honeycomb structures [4, 5], as well as structures based on thrice periodic minimal surfaces (TPMS). Such surfaces have been the subject of much research [6–11] due to their excellent strength properties at high porosity values, which enables the design of lightweight yet strong structures. One of the main advantages of porous materials having a TPMS structure consists in the possibility of varying their properties (thermophysical, mechanical, etc.) by changing the characteristic geometrical parameters of the minimal surfaces [8].

Despite the above-mentioned advantages, a number of difficulties can be encountered when using porous TPMS-materials in the design of heat engineering equipment due to the absence of mathematical models for describing the transfer processes (heat, mass, momentum). Thus, due to the large expenditures of time and resources entailed in the manufacture of prototypes and full-scale experiments to obtain the desired information, the mathematical modeling of heat and mass transfer in TPMS-based porous media becomes an important task.

Both numerical and analytical methods can be used to solve heat and mass transfer problems [12–14]. Analytical methods, such as the method of separation of variables and the method of integral transformations, may be used to obtain exact solutions in particular cases. However, such methods are applicable only to solving a limited number of problems within certain constraints and conditions. When solving complex problems (with nonlinear properties or complex geometry), numerical methods (finite difference method, finite element method, etc.) are often used in practice [15, 16] to find approximate solutions with a given accuracy.

In the present work, a one-dimensional problem of heat transfer of an incompressible fluid flowing in a porous medium having a gyroid-type TPMS structure is solved using the finite difference method.

TASK STATEMENT

The transfer of heat in a porous medium having a gyroid-based structure at a low fluid velocity characteristic of filtration flows is as described in [17, 18]. The schematic of the problem is shown in Fig. 1. Figure 1a shows a unit cell of the gyroid whose characteristic geometric parameters are expressed in terms of the wall thickness δ and length of the cube edge *a* in which the cell is inscribed. When varying these parameters, the porosity of the material changes to affect both the hydrodynamic characteristics of the flow and the intensity of heat transfer.

The porous medium through which the fluid flows is shown in Fig. 1b. The TPMS lattice is formed by copying the unit cell (Fig. 1a) in orthogonal directions of the Cartesian coordinate system with a period equal to *a*.

To describe the heat transfer in fluid flow through a porous medium, we will use the two-temperature model proposed by Wakao and Kagei [19]. The use of separate energy equations to describe each component of the porous medium (solid and fluid) for each phase two-temperature model provides a basis for a more accurate description of the heat transfer process, especially in the absence of local equilibrium.

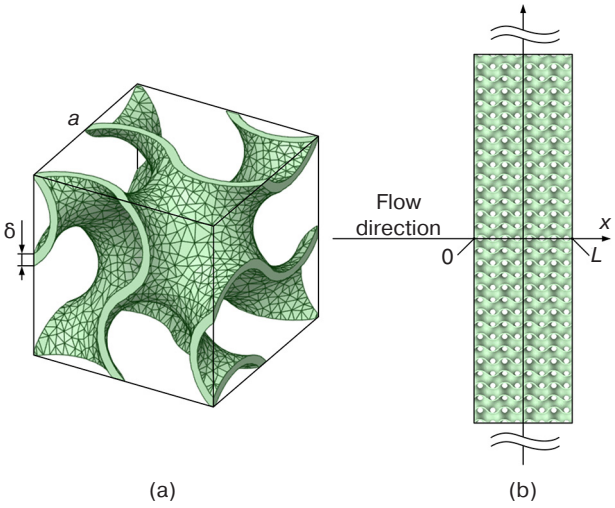


Fig. 1. Task diagram:
(a) gyroid unit cell; (b) porous TPMS medium.
x is a coordinate, *L* is a porous material thickness

The mathematical formulation of the problem, including basic differential equations, as well as initial and boundary conditions, is as follows:

$$(1-\phi)(\rho c_{\mathrm{p}})_{\mathrm{s}}\frac{\partial T_s(x,t)}{\partial t}=$$

$$=\lambda_{\mathrm{eff.s}}\frac{\partial^2 T_{\mathrm{s}}(x,t)}{\partial x^2}+\alpha_{\mathrm{s.f}}(T_{\mathrm{f}}(x,t)-T_{\mathrm{s}}(x,t)), \quad (1)$$

$$(0<x<L, t>0);$$

$$\phi(\rho c_p)_f\left[\frac{\partial T_f(x,t)}{\partial t}+u\frac{\partial T_f(x,t)}{\partial x}\right]=$$
$$=\lambda_{eff.f}\frac{\partial^2 T_f(x,t)}{\partial x^2}+\alpha_{s.f}(T_s(x,t)-T_f(x,t)), \quad (2)$$
$$(0<x<L, t>0);$$

$$T_s(x,0)=T_f(x,0)=T_0; \quad (3)$$

$$-\lambda_s\frac{\partial T_s(x,t)}{\partial x}\bigg|_{x=0}=\alpha_1(T_1-T_s(0,t)); \quad (4)$$

$$\lambda_s\frac{\partial T_s(x,t)}{\partial x}\bigg|_{x=L}=\alpha_2(T_2-T_s(L,t)); \quad (5)$$

$$T_f(0,t)=T_1; \quad (6)$$

$$\frac{\partial T_f(x,t)}{\partial x}\bigg|_{x=L}=0, \quad (7)$$

where t is the time; $T_s(x, t)$ and $T_f(x, t)$ are temperature functions of the solid frame and fluid, respectively; ϕ is the porosity; u is the fluid flow velocity; $(\rho c_p)_s$ and $(\rho c_p)_f$ are density and heat capacity of solid material and fluid, respectively; $\lambda_{eff.s}$ and $\lambda_{eff.f}$ are effective thermal conductivity of solid frame and fluid; T_0 is the initial temperature; T_1 and T_2 are temperatures of surrounding media before and after porous zone; $\alpha_{s.f}$ is the heat transfer coefficient between the fluid and the solid frame; α_1 is the heat transfer coefficient at the boundary between the solid frame and the environment at the point $x = 0$; α_2 is the heat transfer coefficient at the boundary between the solid frame and the environment at the point $x = L$.

Since the heat transfer problem is considered separately from the problem of mass transfer in a porous medium (in some particular cases, the flow dynamics may have little influence on the thermal effects), the equations of conservation of mass and momentum are not included in the system of equations. In addition, the properties of fluid and solid will change insignificantly in the range of considered temperatures. For this reason, the functions to describe the temperature dependence of the properties of the corresponding phases are also not included in the problem formulation.

The same temperature T_0 is taken as the initial condition for both solid and fluid. The boundary condition of the 3rd kind is set at the boundaries of the solid frame (at the points $x = 0$ and $x = L$). The fluid temperature at the point $x = 0$ has a constant value that corresponds to the ambient temperature at the entrance to the porous zone. Since there is no explicit data on the fluid temperature at the point $x = L$, we assume that the heat flux at the boundary is equal to zero.

NUMERICAL SOLUTION

The solution of the problem (1)–(7) is carried out by the finite difference method. For this purpose, let us introduce a spatiotemporal grid, where N is the number of steps in coordinate, and M is the number of steps in time:

$$x^i=i\Delta x, \quad i\in[0,N],$$
$$t^n=n\Delta t, \quad n\in[0,M],$$

where Δx, Δt are coordinate and time steps, respectively.

The finite-difference approximation of differential equations (1), (2) and boundary conditions (3)–(7) is as follows:

$$(1-\phi)(\rho c_p)_s\left[\frac{T_s^{n+1,i}-T_s^{n,i}}{\Delta t}\right]=$$
$$=\lambda_{eff.s}\left[\frac{T_s^{n,i-1}-2T_s^{n,i}+T_s^{n,i+1}}{\Delta x^2}\right]+\alpha_{s.f}(T_f^{n,i}-T_s^{n,i}); \quad (8)$$

$$\phi(\rho c_p)_f\left[\frac{T_f^{n+1,i}-T_f^{n,i}}{\Delta t}+u\frac{T_f^{n,i+1}-T_f^{n,i}}{\Delta x}\right]=$$
$$=\lambda_{eff.f}\left[\frac{T_f^{n,i-1}-2T_f^{n,i}+T_f^{n,i+1}}{\Delta x^2}\right]+\alpha_{s.f}(T_s^{n,i}-T_f^{n,i}); \quad (9)$$

$$T_s^{0,i}=T_f^{0,i}=T_0; \quad (10)$$

$$-\lambda_s\frac{T_s^{n,1}-T_s^{n,0}}{\Delta x}=\alpha_1(T_1-T_s^{n,0}); \quad (11)$$

$$\lambda_s\frac{T_s^{n,N}-T_s^{n,N-1}}{\Delta x}=\alpha_2(T_2-T_s^{n,N}); \quad (12)$$

$$T_f^{0,i}=T_1; \quad (13)$$

$$\frac{T_f^{n,N}-T_f^{n,N-1}}{\Delta x}=0. \quad (14)$$

The task is solved according to an explicit scheme to obtain the values of temperatures T_f and T_s at each grid node at each time step. The explicit scheme was chosen due to its simple implementation and sufficient accuracy for the problem under consideration. At the same time, the stability of the solution is ensured by selecting the time step Δt depending on Δx in accordance with the Courant criterion.

To implement the algorithm for solving the problem by the finite difference method, a software program has been developed in the Python programming language. The graphical interface of the program, which is based on the Tkinter library is shown in Fig. 2.

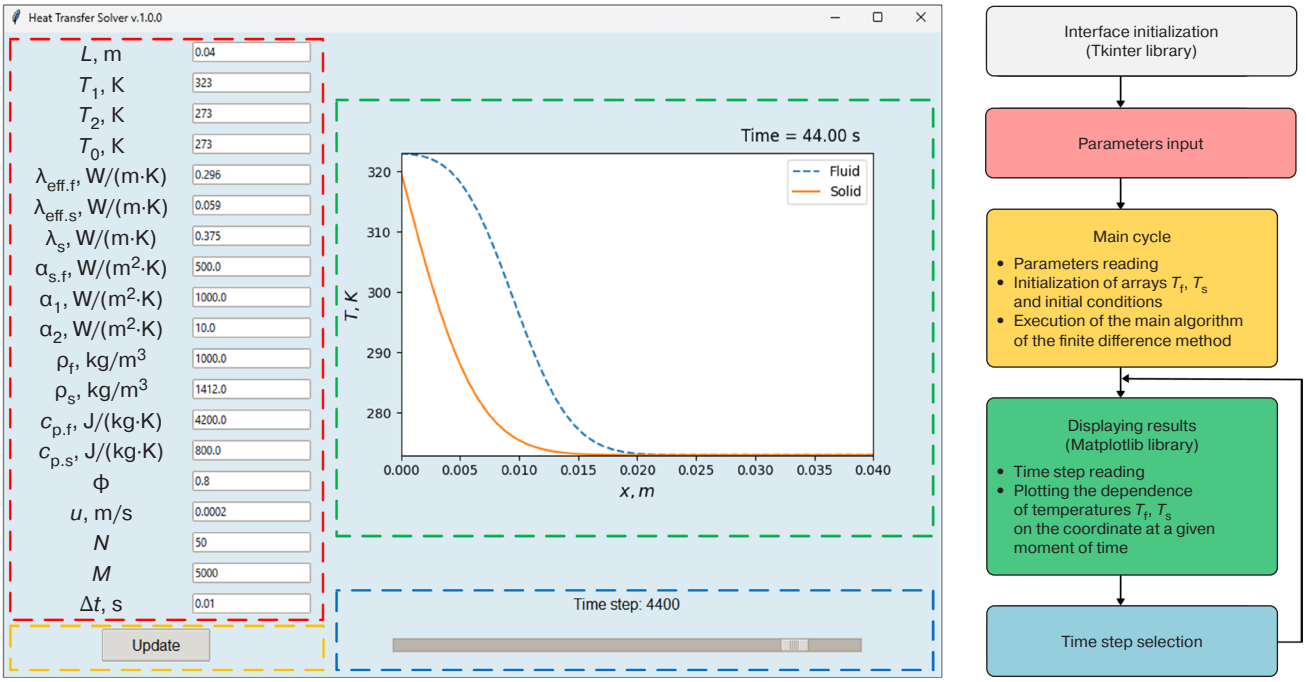


Fig. 2. Interface and algorithm of the developed program

```
for n in range(1, M):
    for i in range(1, N - 1):
        T_s[n, i] = ((1 / ((1 - phi) * p_s * co_s)) *
                     (((T_s[n - 1, i - 1] - 2 * T_s[n - 1, i] + T_s[n - 1, i + 1]) / dx ** 2) * k_es +
                     h_sf * (T_f[n - 1, i] - T_s[n - 1, i])) * dt + T_s[n - 1, i])

        T_f[n, i] = (((k_ef * ((T_f[n - 1, i - 1] - 2 * T_f[n - 1, i] + T_f[n - 1, i + 1]) / dx ** 2) +
                     h_sf * (T_s[n - 1, i] - T_f[n - 1, i])) / (phi * p_f * co_f)) - u * (T_f[n - 1, i] -
                     T_f[n - 1, i - 1]) / dx) * dt + T_f[n - 1, i]

    T_s[n, 0] = (h_ext1 * dx * T_ext1 + k_s * T_s[n, 1]) / (k_s + h_ext1 * dx)

    T_s[n, -1] = (k_s * T_s[n, -2] + h_ext2 * dx * T_ext2) / (k_s + h_ext2 * dx)

    T_f[n, -1] = T_f[n, -2]
```

Fig. 3. Finite-difference diagram in the Python programming language

The following stages can be distinguished in the main cycle of the program:

1. Input of the main parameters of the problem through special text fields. The program functionality allows both the thermophysical properties of the studied materials and the size of the spatiotemporal grid to be set manually.
2. Reading of parameters, initialization of arrays for T_f, T_s and execution of the main cycle of the finite difference method. The finite-difference scheme for the basic differential equations (8), (9) and boundary conditions (11), (12), (14) in the Python programming language is presented in Fig. 3.
3. Visualization of the results is provided by plotting the diagram of temperature change of the solid frame and fluid within the considered area. The Matplotlib library was used to display an interactive diagram of temperature change in a porous medium. The graph display can be controlled by means of a slider to change the time step and update the temperature field plots. This provides a means of studying the dynamics of the heat transfer process and the influence of various parameters on the temperature distribution.

The program was tested for correctness by comparing numerical solutions with known analytical solutions for particular cases of the problem. The source code as well as the executable file of the program are presented on the Mendeley Data[1] repository [20].

[1] https://data.mendeley.com/datasets/kcn33tr7sb/2. Accessed July 23, 2025.

RESULTS

Using the developed program, the problem of calculating heat transfer in a gyroid-based porous medium has been solved. Graphical dependences of the temperature change of the solid frame and the fluid were obtained. In particular, a graph of the dependence of the temperature of the fluid and the solid frame on the coordinate at different moments of time is presented (Fig. 4). This graph was obtained with the following constants and values of thermophysical properties of the solid frame and fluid, as well as boundary and initial conditions: $c_{p.s}$ = 800 J/(kg·K), ρ_s = 1412 kg/m^3, λ_s = 0.375 W/(m·K), $c_{p.f}$ = 4200 J/(kg·K), ρ_f = 1000 kg/m^3, λ_f = 0.6 W/(m·K), $\lambda_{eff.f}$ = 0.296 W/(m·K), $\lambda_{eff.s}$ = 0.059 W/(m·K), $\phi = 0.8$, L = 0.04 m, T_0 = 273 K, T_1 = 323 K, T_2 = 273 K, α_1 = 1000 W/(m^2·K), α_2 = 10 W/(m^2·K), $\alpha_{s.f}$ = 500 W/(m^2·K), u = 0.0002 m/s. The coefficients of effective thermal conductivity of solid and fluid phases in porous material with gyroid-based structure were determined according to the methodology presented in [21].

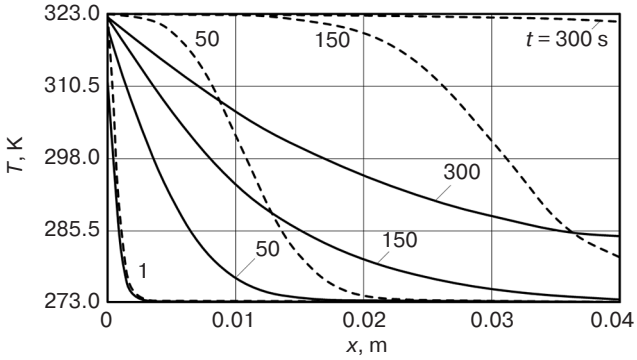


Fig. 4. Diagram of temperature distribution in a porous medium:
------ solid framework; – – – – fluid

From the diagram in Fig. 4 it can be seen that the temperatures of the fluid and the solid frame differ significantly over the entire range of the spatial coordinate under the given conditions. This emphasizes the importance of using a two-temperature model to accurately describe the temperature distribution in porous media.

The diagram in Fig. 5 demonstrates the change in temperature at different points of the porous material, namely, at the beginning, in the middle and at the end of the region under consideration. From the analysis of this graph, it can be concluded that under the given heat exchange conditions, the fluid temperature reaches the maximum value specified by the boundary condition (T_1 = 323 K). At t > 300 s, the temperature curves of the fluid and the solid framework begin to converge, which indicates a decrease in the absolute value of the temperature gradient and an approach to thermodynamic equilibrium in the system.

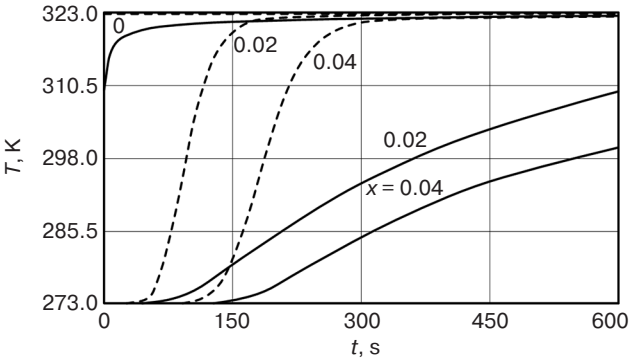


Fig. 5. Diagram of temperature change in a porous medium over time:
------ solid framework; – – – – fluid

Figure 6 shows a diagram of the heat flux distribution determined in accordance with the Newton–Richmann law ($|q| = \alpha_{s.f}(T_f - T_s)$) in a porous material along the coordinate. The pronounced peak in the heat flux distribution near the left boundary of the region at the early stages of the process ($t \approx 10$ s) is associated with a large absolute value of the temperature gradient caused by the adoption of the first-order boundary condition at the point x = 0. As time increases (10 s < t < 200 s), an increase in the heat flux is observed with a shift in the peak value to the right boundary of the region. The highest efficiency of the process observed at time $t \approx 200$ s is followed by a gradual decrease in the heat exchange intensity; at this point, the fluid temperature has the maximum possible value in the entire range of x (Fig. 3).

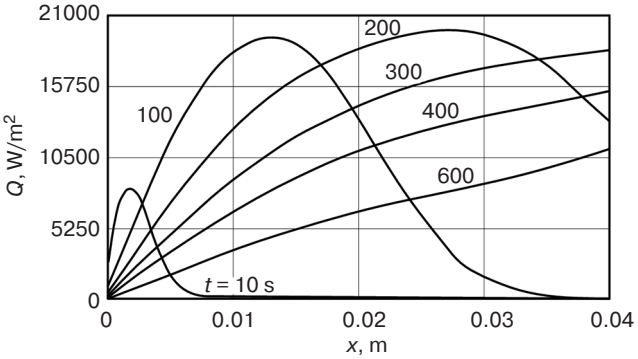


Fig. 6. Heat flow distribution diagram

CONCLUSIONS

We have presented a study of heat transfer in a porous medium with an ordered macrostructure based on a gyroid during filtration flow. For mathematical modeling of the process under study, a Wakao and Kagei two-temperature model was used.

In order to obtain the numerical solution of the heat transfer problem in a porous medium (1)–(7) using the finite difference method, a program in the Python programming language was developed. The functionality of the developed program provides for the construction of a graph of the temperature of a solid frame and fluid. The visualization of changes in temperature curves enabled by the built-in time step slider is important for studying the dynamics of the heat transfer process. As compared with direct modeling methods (for example, the finite element method in *ANSYS*[2]), the program represents a valuable means for quickly obtaining a solution to the problem, which is achieved due to the lack of need to construct computational geometry and a computational grid.

Porous materials having a TPMS-based (e.g. gyroid) structure can be used in the design of heat exchangers, catalysts and other devices where it is important to minimize the mass of the structure without compromising its strength and efficiency. The developed program can be used to model heat transfer in porous TPMS as a means for quickly predicting their behavior under real operating conditions. This has potential utility in solving a number of engineering problems in the energy, chemical industry, mechanical engineering, and other industries.

The results of this work can become the basis for further scientific research in the field of modeling heat and mass transfer in porous media. In the future, it is possible to expand the model and the program by taking into account the equation of fluid motion in a porous medium to determine both the velocity profile and pressure losses. An additional promising direction may consist in the development of similar models for other types of minimal surfaces to select structures having the desired heat-transfer-, strength-, and mass characteristics depending on the engineering tasks set.

ACKNOWLEDGMENTS

The study was supported by the Russian Science Foundation, grant No. 23-79-10044, https://rscf.ru/project/23-79-10044/.

Authors' contributions

A.I. Popov—researching, writing the manuscript, developing of the program, visualization, and editing the article.

A.V. Eremin—task statement, funding acquisition, and editing the article.

REFERENCES

1. Popov I.A., Gortyshov Yu.F., Olimpiev V.V. Industrial application of heat transfer enhancement: The modern state of the problem (a Review). *Therm. Eng*. 2012;59(1):1–12. https://doi.org/10.1134/S0040601512010119
[Original Russian Text: Popov I.A., Gortyshov Yu.F., Olimpiev V.V. Industrial application of heat transfer enhancement: The modern state of the problem (a Review). *Teploenergetika.* 2012;1:3–14 (in Russ.).]
2. Solovev S.A., Soloveva O.V., Shakurova R.Z., Golubev Ya.P. Overview of the application of open cell foam heat exchangers. *Izvestiya vysshikh uchebnykh zavedenii. PROBLEMY ENERGETIKI = Power Engineering: Research, Equipment, Technology*. 2024;26(1):165–194 (in Russ.). https://doi.org/10.30724/1998-9903-2024-26-1-165-194
3. Kugatov P.V. Use of porous carbon materials as carriers for catalysts. *Bashkirskii khimicheskii zhurnal = Bashkir Chemical Journal*. 2011;18(1):98–105 (in Russ.).
4. Testoedov N.A., Nagovitsin V.N., Permyakov M.Yu. Spacecraft application of three layer honeycomb structures. *Sibirskii aerokosmicheskii zhurnal = Siberian Aerospace Journal*. 2016;17(1):200–211 (in Russ.).
5. Solovev S.A., Soloveva O.V., Shakurova R.Z. Review of modern ceramic cellular materials and composites used in heat engineering. *Izvestiya vysshikh uchebnykh zavedenii. PROBLEMY ENERGETIKI = Power Engineering: Research, Equipment, Technology.* 2023;25(1):82–104 (in Russ.). https://doi.org/10.30724/1998-9903-2023-25-1-82-104
6. Bragin D.M., Eremin A.V., Popov A.I., Shulga A.S. Method to determine effective thermal conductivity coefficient of porous material based on minimum surface Schoen's I-WP(R) type. *Vestnik IGEU*. 2023;2:61–68 (in Russ.). https://doi.org/10.17588/2072-2672.2023.2.061-068
7. Zinina S.A., Popov A.I., Eremin A.V. Numerical solution of the nonlinear problem of thermal conductivity in a porous plate with an ordered macrostructure. *Vestnik TvGU. Seriya: Prikladnaya matematika = Herald of Tver State University. Ser.: Appl. Math*. 2024;1:53–67. https://doi.org/10.26456/vtpmk702
8. Popov A.I. Development of thermal insulation with ordered structure based on Neovius TPMS. *Vestnik IGEU*. 2022;6:58–68 (in Russ.).
9. Al-Ketan O., Abu Al-Rub R.K. Multifunctional mechanical metamaterials based on triply periodic minimal surface lattices. *Adv. Eng. Mater*. 2019;21(10):1900524. https://doi.org/10.1002/adem.201900524
10. Schoen A.H. Reflections concerning triply-periodic minimal surfaces. *Interface Focus*. 2012;2(5):658–668. https://doi.org/10.1098/rsfs.2012.0023

[2] https://www.cadfem-cis.ru/ (in Russ.). Accessed July 21, 2025.

11. Abueidda D.W., Bakir M., Al-Rub R.K.A., Bergström J.S., Sobh N.A., Jasiuk I. Mechanical properties of 3D printed polymeric cellular materials with triply periodic minimal surface architectures. *Materials & Design*. 2017;122(9):255–267. https://doi.org/10.1016/j.matdes.2017.03.018
12. Kartashov E.M. Mathematical models of heat conduction with two-phase lag. *J. Eng. Phys. Thermophys*. 2016;89(2): 346–356. https://doi.org/10.1007/s10891-016-1385-9
[Original Russian Text: Kartashov E.M. Mathematical models of heat conduction with two-phase lag. *Inzhenerno-fizicheskii zhurnal*. 2016;89(2):338–349 (in Russ.).]
13. Kartashov E.M., Krotov G.S. Analytical solution of single-phase Stefan problem. *Math. Models Comput. Simul*. 2009;1(2):180–188. https://doi.org/10.1134/S2070048209020021
[Original Russian Text: Kartashov E.M., Krotov G.S. Analytical solution of single-phase Stefan problem. *Matematicheskoe modelirovanie*. 2008;20(3):77–86 (in Russ.).]
14. Kartashov E.M. Analytical approaches to the analysis of unsteady heat conduction for partially bounded regions. *High Temp.* 2020;58(3):377–385. https://doi.org/10.1134/S0018151X20030086
[Original Russian Text: Kartashov E.M. Analytical approaches to the analysis of unsteady heat conduction for partially bounded regions. *Teplofizika vysokikh temperatur.* 2020;58(3):402–411 (in Russ.). https://doi.org/10.31857/S0040364420030084]
15. Korenchenko A.E., Zhukova A.A. Evaporation of a liquid sessile droplet subjected to forced convection. *Russ. Technol. J*. 2021;9(5):57–66 https://doi.org/10.32362/2500-316X-2021-9-5-57-66
16. Glinskiy I.A., Zenchenko N.V., Maltsev P.P. Thermal modelling of terahertz Quantum-cascade laser based on nanoheterostructures GaAs/AlGaAs. *Rossiiskii Tekhnologicheskii Zhurnal*. 2016;4(3):27–36 (in Russ.). https://doi.org/10.32362/2500-316X-2016-4-3-27-36
17. Hayashi K., Kishida R., Tsuchiya A., Ishikawa K. Superiority of triply periodic minimal surface gyroid structure to strut-based grid structure in both strength and bone regeneration. *ACS Appl. Mater. Interfaces*. 2023;15(29):34570–34577. https://doi.org/10.1021/acsami.3c06263
18. Chouhan G., Bala Murali G. Designs, advancements, and applications of three-dimensional printed gyroid structures: A review. In: *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*. 2024;238(2):965–987. https://doi.org/10.1177/09544089231160030
19. Wakao N., Kagei S. *Heat and Mass Transfer in Packed Beds*. Taylor & Francis; 1982. 364 p.
20. Popov A.I. Heat Transfer Solver. V. 1. *Mendeley Data*. 2024. https://www.doi.org/10.17632/kcn33tr7sb.1
21. Bragin D.M., Popov A.I., Eremin A.V. The thermal conductivity properties of porous materials based on TPMS. *Int. J. Heat Mass Transfer*. 2024;231:125863. https://doi.org/10.1016/j.ijheatmasstransfer.2024.125863

СПИСОК ЛИТЕРАТУРЫ

1. Попов И.А., Гортышов Ю.Ф., Олимпиев В.В. Промышленное применение интенсификации теплообмена – современное состояние проблемы (Обзор). *Теплоэнергетика*. 2012;1:3–14.
2. Соловьев С.А., Соловьева О.В., Шакурова Р.З., Голубев Я.П. Обзор применения высокопористых ячеистых теплообменников. *Известия высших учебных заведений. ПРОБЛЕМЫ ЭНЕРГЕТИКИ*. 2024;26(1):165–194. https://doi.org/10.30724/1998-9903-2024-26-1-165-194
3. Кугатов П.В. Использование пористых углеродных материалов в качестве носителей для катализаторов. *Башкирский химический журнал*. 2011;18(1):98–105.
4. Тестоедов Н.А., Наговицин В.Н., Пермяков М.Ю. Применение трехслойных сотовых конструкций в космических аппаратах. *Сибирский аэрокосмический журнал*. 2016;17(1):200–211.
5. Соловьева О.В., Соловьев С.А., Шакурова Р.З. Обзор современных керамических ячеистых материалов и композитов, применяемых в теплотехнике. *Известия высших учебных заведений. ПРОБЛЕМЫ ЭНЕРГЕТИКИ*. 2023;25(1): 82–104. https://doi.org/10.30724/1998-9903-2023-25-1-82-104
6. Брагин Д.М., Еремин А.В., Попов А.И., Шульга А.С. Метод определения коэффициента эффективной теплопроводности пористого материала на основе минимальной поверхности типа Schoen's I-WP(R). *Вестник ИГЭУ*. 2023;2: 61–68. https://doi.org/10.17588/2072-2672.2023.2.061-068
7. Зинина С.А., Попов А.И., Еремин А.В. Численное решение нелинейной задачи теплопроводности в пористой пластине с упорядоченной макроструктурой. *Вестник ТвГУ. Серия: Прикладная математика*. 2024;1:53–67. https://doi.org/10.26456/vtpmk702
8. Попов А.И. Разработка тепловой изоляции с упорядоченной структурой, основанной на ТПМП Неовиуса. *Вестник ИГЭУ*. 2022;6:58–68.
9. Al-Ketan O., Abu Al-Rub R.K. Multifunctional mechanical metamaterials based on triply periodic minimal surface lattices. *Adv. Eng. Mater*. 2019;21(10):1900524. https://doi.org/10.1002/adem.201900524
10. Schoen A.H. Reflections concerning triply-periodic minimal surfaces. *Interface Focus*. 2012;2(5):658–668. https://doi.org/10.1098/rsfs.2012.0023
11. Abueidda D.W., Bakir M., Al-Rub R.K.A., Bergström J.S., Sobh N.A., Jasiuk I. Mechanical properties of 3D printed polymeric cellular materials with triply periodic minimal surface architectures. *Materials & Design*. 2017;122(9):255–267. https://doi.org/10.1016/j.matdes.2017.03.018

12. Карташов Э.М. Математические модели теплопроводности с двухфазным запаздыванием. *Инженерно-физический журнал*. 2016;89(2):338–349.
13. Карташов Э.М., Кротов Г.С. Аналитическое решение однофазной задачи Стефана. *Математическое моделирование*. 2008;20(3):77–86.
14. Карташов Э.М. Аналитические подходы к исследованиям нестационарной теплопроводности для частично ограниченных областей. *Теплофизика высоких температур*. 2020;58(3):402–411. https://doi.org/10.31857/S0040364420030084
15. Коренченко А.Е., Жукова А.А. Испарение жидкой лежащей капли в условиях вынужденной конвекции. *Russ. Technol. J*. 2021;9(5):57–66. https://doi.org/10.32362/2500-316X-2021-9-5-57-66
16. Глинский И.А., Зенченко Н.В., Мальцев П.П. Тепловое моделирование терагерцового квантового-каскадного лазера на основе наногетероструктуры GaAs/AlGaAs. *Rossiiskii Tekhnologicheskii Zhurnal*. 2016;4(3):27–36. https://doi.org/10.32362/2500-316X-2016-4-3-27-36
17. Hayashi K., Kishida R., Tsuchiya A., Ishikawa K. Superiority of triply periodic minimal surface gyroid structure to strut-based grid structure in both strength and bone regeneration. *ACS Appl. Mater. Interfaces*. 2023;15(29):34570–34577. https://doi.org/10.1021/acsami.3c06263
18. Chouhan G., Bala Murali G. Designs, advancements, and applications of three-dimensional printed gyroid structures: A review. In: *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*. 2024;238(2):965–987. https://doi.org/10.1177/09544089231160030
19. Wakao N., Kagei S. *Heat and Mass Transfer in Packed Beds*. Taylor & Francis; 1982. 364 p.
20. Popov A.I. Heat Transfer Solver. V. 1. *Mendeley Data*. 2024. https://www.doi.org/10.17632/kcn33tr7sb.1
21. Bragin D.M., Popov A.I., Eremin A.V. The thermal conductivity properties of porous materials based on TPMS. *Int. J. Heat Mass Transfer*. 2024;231:125863. https://doi.org/10.1016/j.ijheatmasstransfer.2024.125863

About the Authors

Andrey I. Popov, Cand. Sci. (Eng.), Senior Lecturer, Department of Industrial Heat Power Engineering, Samara State Technical University (244, Molodogvardeyskaya ul., Samara, 443100 Russia). E-mail: popov.ai@samgtu.ru. Scopus Author ID 57216363622, RSCI SPIN-code 5560-6869, https://orcid.org/0000-0001-5014-8167

Anton V. Eremin, Dr. Sci. (Eng.), Associate Professor, Head of Department of Industrial Heat Power Engineering, Samara State Technical University (244, Molodogvardeyskaya ul., Samara, 443100 Russia). E-mail: a.v.eremin@list.ru. Scopus Author ID 56395547000, ResearcherID D-6936-2014, RSCI SPIN-code 3892-0775, https://orcid.org/0000-0002-2614-6329

Об авторах

Попов Андрей Игоревич, к.т.н., старший преподаватель, кафедра «Промышленная теплоэнергетика», ФГБОУ ВО «Самарский государственный технический университет» (443100, Россия, Самара, ул. Молодогвардейская, д. 244). E-mail: popov.ai@samgtu.ru. Scopus Author ID 57216363622, SPIN-код РИНЦ 5560-6869, https://orcid.org/0000-0001-5014-8167

Еремин Антон Владимирович, д.т.н., доцент, проректор по интеграционным проектам, ФГБОУ ВО «Самарский государственный технический университет» (443100, Россия, Самара, ул. Молодогвардейская, д. 244). E-mail: a.v.eremin@list.ru. Scopus Author ID 56395547000, ResearcherID D-6936-2014, SPIN-код РИНЦ 3892-0775, https://orcid.org/0000-0002-2614-6329

Translated from Russian into English by L. Bychkova
Edited for English language and spelling by Thomas A. Beavitt

Mathematical modeling

Математическое моделирование

UDC 621.372.8
https://doi.org/10.32362/2500-316X-2025-13-5-104-118
EDN ZTAAYP

RESEARCH ARTICLE

Modeling of thermophysical processes in an oil reservoir during heating in a stopped well

Sergey E. Savotchenko [1, 2, @], **Vasily A. Zakharov** [2]

[1] *MIREA – Russian Technological University, Moscow, 119454 Russia*
[2] *Sergo Ordzhonikidze Russian State University for Geological Prospecting, Moscow, 117997 Russia*
[@] *Corresponding author, e-mail: savotchenkose@mail.ru*

• **Submitted:** 15.07.2024 • **Revised:** 08.12.2024 • **Accepted:** 23.07.2025

Abstract

Objectives. An important and urgent task of the oil producing industry is the identification of patterns of thermophysical processes in reservoirs. One approach to improving the efficiency of oil recovery in conditions of hard-to-recover reserves involves thermal action on the reservoir. The construction of mathematical models for describing such processes to optimize production technologies is based on the formation of nonstationary heat flows in the reservoir when a stopped well is heated. The application of mathematical modeling methods considered in the work forms a basis for calculating the distribution dependencies of nonstationary fields of thermophysical characteristics in the reservoir when heating the well to its parameters and the properties of the environments.

Methods. The work is based on heat- and mass-transfer theory along with mathematical physics, analytical and numerical methods, as well as algorithms, computer modeling approaches, and the development of applications using modern programming languages and their libraries.

Results. A formation saturated with oil, water, and a steam–gas mixture is theoretically described. A closed system of heat and mass transfer equations is obtained taking into account diffusion-droplet and heat flows and phase transformations. A formulated mathematical statement of the model comprises an initial–boundary value problem for equations relating the temperature, saturation, and pressure of the components of the saturating fluid in the formation. Numerical algorithms for solving are developed and their software implementation carried out. An application developed for computer implementation of the model provides convenient visualization of the calculation results consisting of several components (modules). Numerical experiments were carried out using the developed software to study how various factors, such as the properties of the formation sketch and the saturating liquid phase and heater characteristics, affect the thermophysical processes in the formation.

Conclusions. The developed model can be used to clearly describe nonstationary distributions of thermophysical characteristics formed by thermal and diffusion-droplet flows in the reservoir during heating of a shut-up well. The obtained results expand current understandings of the regularities of thermophysical processes and the properties of the saturating phase in the reservoir under thermal influence.

Keywords: thermophysical processes, heat transfer, mass transfer, heat flow, diffusion-droplet flow, heat equation, heat transfer equation, thermal conductivity, thermal impact on the formation, oil well heating

For citation: Savotchenko S.E., Zakharov V.A. Modeling of thermophysical processes in an oil reservoir during heating in a stopped well. *Russian Technological Journal.* 2025;13(5):104–118. https://doi.org/10.32362/2500-316X-2025-13-5-104-118, https://www.elibrary.ru/ZTAAYP

Financial disclosure: The authors have no financial or proprietary interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Моделирование теплофизических процессов в нефтяном пласте при прогреве в остановленной скважине

С.Е. Савотченко [1, 2, @], **В.А. Захаров** [2]

[1] *МИРЭА – Российский технологический университет, Москва, 119454 Россия*
[2] *Российский государственный геологоразведочный университет им. Серго Орджоникидзе, Москва, 117997 Россия*
[@] *Автор для переписки, e-mail: savotchenkose@mail.ru*

• **Поступила:** 15.07.2024 • **Доработана:** 08.12.2024 • **Принята к опубликованию:** 23.07.2025

Резюме

Цели. Выявление закономерностей теплофизических процессов в пластах является важной и актуальной задачей нефтедобывающей отрасли. Одним из способов повышения эффективности нефтеотдачи в условиях трудноизвлекаемых запасов является тепловое воздействие на пласт. При нагреве остановленной скважины в пласте формируются нестационарные тепловые потоки, поэтому в вопросах оптимизации добывающих технологий таких процессов широко применяется построение адекватных математических моделей. Цель работы – развитие возможностей применения методов математического моделирования и установление на их основе зависимостей распределения нестационарных полей теплофизических характеристик в пласте при нагревании скважины от ее параметров и свойств сред.

Методы. Использованы теория тепло- и массопереноса, методы математической физики, аналитические и численные методы, алгоритмы, методы компьютерного моделирования и разработки приложений, современные языки программирования и их библиотеки.

Результаты. Проведено теоретическое описание пласта, насыщенного нефтью, водой и парогазовой смесью. Получена замкнутая система уравнений тепло- и массопереноса при учете диффузионно-капельных и тепловых потоков и фазовых превращений. Сформулирована математическая постановка модели, представляющая собой начально-краевую задачу для уравнений, связывающих температуру, насыщенность и давление компонентов насыщающей жидкости в пласте. Разработаны численные алгоритмы решения такой задачи и проведена их программная реализация. Разработано приложение для компьютерной реализации модели с удобной визуализацией результатов расчетов, состоящей из нескольких компонентов (модулей). С использованием разработанного программного обеспечения проведены численные эксперименты для изучения того, как различные факторы, такие как свойства скетча пласта и насыщающей жидкой фазы, характеристики нагревателя, влияют на теплофизические процессы в пласте.

Выводы. Разработанная модель позволяет наглядно описать нестационарные распределения теплофизических характеристик, формируемых тепловым и диффузионно-капельным потоками в пласте в процессе прогрева остановленной скважины. Полученные результаты расширяют представления о закономерностях теплофизических процессов и свойствах насыщающей фазы в пласте при тепловом воздействии.

Ключевые слова: теплофизические процессы, теплоперенос, массоперенос, тепловой поток, диффузионно-капельный поток, уравнение теплопроводности, уравнение теплопереноса, теплопроводность, тепловое воздействие на пласт, прогрев скважины

Для цитирования: Савотченко С.Е., Захаров В.А. Моделирование теплофизических процессов в нефтяном пласте при прогреве в остановленной скважине. *Russian Technological Journal*. 2025;13(5):104–118. https://doi.org/10.32362/2500-316X-2025-13-5-104-118, https://www.elibrary.ru/ZTAAYP

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

The development of mathematical models of thermophysical processes in the reservoir as a means of increasing the efficiency of oil recovery, especially in conditions of hard-to-recover reserves, remains one of the urgent tasks in the oil producing industry [1]. Since thermal stimulation of the reservoir is one of the most efficient means to increase oil recovery under such conditions [2–4], the use of mathematical models to obtain various kinds of information on the heat transfer patterns in the reservoir is an important area of research [5, 6]. Nonstationary heat flows in the reservoir are formed during heating of a shut-in well. When the thermal stimulation ends, the reservoir begins to cool. The need to describe such processes is associated with economic imperative of maximizing oil production [7, 8]. To develop effective strategies, it is necessary to understand the dynamics of changes in temperature, saturation, and pressure in the reservoir [9]. Mathematical modeling is the main tool used in such research due to the need to consider the complex interaction of physical processes [10, 11]. The formulation of mathematical models of thermophysical processes is based on a system of heat transfer equations used in many thermophysical problems and heat conduction theory [12, 13].

Mathematical modeling of thermophysical processes in oil-producing wells and formations is described in many works [14–16]. Such studies have been actively continued in recent years using numerical methods and computer modeling [17, 18]. In particular, Tupysev [19] studied the regularities of non-isothermal gas filtration in a well during the creation of low-temperature gas deposits when the thermobaric conditions of the formation approach the equilibrium conditions of hydrate formation. The described method can be used to predict the exploration and operation of wells during the development of low-temperature gas deposits, as well as to determine the dynamics of possible hydrate formation in the bottomhole zone and the influence of this process on the operation of wells.

Sharafutdinov et al. [20] carried out a numerical study of the temperature field in a multi-layer well during the movement of gasified oil taking the Joule–Thomson effects and heat of degassing into account. The possibility of estimating the position of the boundary of the oil degassing region in the wellbore using the calculated temperature distribution was demonstrated.

Ramazanov and Parshin [21] proposed and analytically investigated a model describing the formation of a temperature field in the reservoir with a combined inflow of formation water and gassed oil to the well. The time of observation of the maximum temperature decrease was shown to be determined by the radius of the degassing zone and the speed of convective heat transfer (which depend on the specific flow rates of oil and water), and this time can be used to estimate such parameters. It was confirmed that, with an increase in water saturation, the influence of the cooling effect of oil degassing regularly decreases.

Ramazanov and Parshin [22] also determined the conditions for observing a non-monotonic change in temperature over time when the temperature at the outlet of a gas–liquid mixture from a porous medium initially decreases for some time and then increases.

Shagapov and Tazetdinov [23] propose a mathematical model describing the formation of a temperature field in a radially symmetric heated formation with high-viscosity oil through a horizontal well. The authors demonstrated the possibility of further exploitation of the well for the production of oil with reduced viscosity. A mathematical statement based on a system of equations is used to describe the heat transfer process for assessing the characteristic ranges of penetration of filtration and temperature waves for the time periods under consideration.

Ramazanov et al. [24] presented an algorithm and software package for calculating and modeling the process of oil displacement during hot water injection to obtain temperature fields in the formation at different points in time, as well as to determine the oil recovery of the formation taking into account the influence of the thermal characteristics of liquid-saturated rocks.

Gil'manov et al. [25] propose a mathematical model of cyclic–steam treatment of a formation taking into account the mass fraction of steam in the coolant

and the equation of state for water. The model is based on the use of heat balance relations at each stage of cyclic steam treatment. The described model is used to determine the steam temperature in the productive interval along with the initial formation temperature. Here, the heat flow calculations are based on the data of short-term dynamic temperature studies, while the oil consumption was determined using the Dupuit formula for a zonally heterogeneous formation. The model was also used to determine the optimal times of the stages of cyclic–steam treatment and the maximum cumulative oil production. The optimal time for pumping the coolant into the formation and the well holding time for steam condensation were shown to increase with an increase in the thickness of the formation, the coolant consumption, and the mass fraction of steam in it.

The importance of the ongoing studies of thermophysical processes in formations and wells is due to the need to understand their patterns for developing technologies based on the thermal impact on the formation aimed at increasing the efficiency of oil extraction. Numerical methods and computer modeling are widely used to describe nonstationary distributions of the temperature field and pressure in formations and wells. However, despite the significant development of this problem already undertaken, a number of issues require more detailed study by methods of mathematical modeling.

The present work presents the results of the development and computer implementation of a model that describes the nonstationary process of heat flow distribution in a reservoir under constant thermal impact in a shut-in well (in which production is temporarily stopped). Based on the developed computer implementation of the model, the patterns and mechanisms of the influence of the thermophysical parameters of the reservoir on the nonstationary distribution of temperature, saturation, and pressure parameters in it are studied.

1. THEORETICAL BASIS OF THE MODEL

In the volume element of a formation saturated with oil, water, and a steam-gas mixture in thermodynamic equilibrium, the temperature T, pressure p, and saturation θ are distributed uniformly at the initial moment of time. If we neglect the compressibility of the formation skeleton and gravitational effects, then the continuity equation for each component of the saturating liquid phase can be written in the form [16, 26–30]:

$$m\frac{\partial(\rho\theta)}{\partial t} = -\mathrm{div}J_\theta + I, \quad (1)$$

where m is the porosity of the average, ρ is the component density; t is time, θ is the volumetric saturation of the mixture component of the element of the formation volume.

$$\theta = \frac{V}{mV_e},$$

where V is the volume of the saturating liquid phase component in the formation volume element; V_e is the volume of the formation element; J_θ is the diffusion-droplet component of the mass density of the flow transferred by the molecular thermal conductivity of the formation volume element; and I is the volumetric power of the source of the liquid component.

In accordance with Fick's law, the diffusion-droplet component of the mass flux density can be represented as

$$J_\theta = -a(\nabla\theta + \delta\nabla T), \quad (2)$$

where a is the coefficient of diffusion-droplet mass transfer of the component of the saturating liquid phase, δ is its thermogradient coefficient, and ∇ is the nabla symbol ($\nabla\theta$ is the saturation gradient and ∇T is the temperature gradient).

The power of the source can be represented as

$$I = \varepsilon\rho m\frac{\partial\theta}{\partial t}, \quad (3)$$

where ε is the phase transition coefficient, representing the ratio of the increment in saturation of a liquid component obtained during a phase transition to the total increment in saturation of the liquid component taking into account diffusion, droplet, and convective processes.

The heat and mass transfer equation, taking into account Eq. (3), can be written in the form [16, 26–30]

$$c\frac{\partial T}{\partial t} = -\mathrm{div}J + \varepsilon q\rho m\frac{\partial\theta}{\partial t}, \quad (4)$$

where c is the specific heat capacity of the volume element of the formation; J is the heat flux density; q is the specific heat of phase transition.

According to Fourier's law, the heat flux density can be represented as

$$J = -\lambda\nabla T, \quad (5)$$

where λ is the effective thermal conductivity of an element of the formation volume.

To the given equations, the equation of state of the mixture should be added [16, 26–29]:

$$\rho\theta = \frac{pM}{zRT}(1+\theta), \quad (6)$$

where M is the molar mass of the mixture; R is the universal gas constant; z is the correction factor that

takes into account the deviation of the vapor–gas mixture from an ideal gas.

Accurate to terms of the first order of smallness [16],

$$\nabla(\rho\theta) \approx \rho_e \beta \frac{\partial p}{\partial t},$$

where β is the coefficient of elasticity of the vapor–gas mixture; ρ_e is the density of the formation skeleton. The continuity equation that closes the system of equations for temperature and saturation can be written in the form:

$$\beta \frac{\partial p}{\partial t} = -\varepsilon \frac{\rho}{\rho_e} \frac{\partial \theta}{\partial t}. \quad (7)$$

The system of Eqs. (1)–(7) forms the theoretical basis for modeling thermophysical processes in the reservoir.

It should be noted that the values of the above-introduced coefficient of diffusion-droplet mass transfer, thermogradient coefficient, and phase transition coefficient are determined from experiments. In the general case, these coefficients may depend on both temperature and saturation. However, in experimentally established ranges of temperatures and saturations for a number of water-saturated specific media (sands, sandstones, clays, ceramics), the indicated coefficients are found to be practically constant.

For example, in the temperature range of 293–423 K, the phase transition coefficient ε is constant [16]. With increasing saturation in the range of 0.3–0.4, its value decreases linearly from 1.0 to 0.3, but with a further increase in saturation, it remains constant. The thermogradient coefficients δ behave completely similarly to the phase transition coefficient depending on temperature and saturation; in particular, its stabilized value in a number of media is $(0.2–0.5) \cdot 10^{-3}$ K^{-1}.

Although the coefficient of diffusion-droplet mass transfer a in the temperature range under consideration increases by approximately 1.5 times in some media (sands, sandstones), in a number of others (ceramics, clays), it is temperature-independent. Its stabilized values in sandy media are $(1.8–8.8) \cdot 10^{-4}$ kg/(m·s), while in ceramics and clays, the corresponding figure is $(4.4–8.9) \cdot 10^{-6}$ kg/(m·s).

The assumption of the constancy of these coefficients in theoretical modeling is valid in the case of low-intensity processes in which the temperature and saturations change insignificantly over short periods of time. For the purposes of constructing a model, an approach is also possible in which the heat and mass transfer processes under study are divided into separate sections in each of which the coefficients under consideration are considered constant.

2. FORMULATION OF THE MODEL AND STATEMENT OF THE PROBLEM

In this paper, we consider the stationary thermal effect of a borehole heater on a reservoir. We consider the reservoir as a continuous, homogeneous, thermally isotropic medium having constant effective values of thermophysical coefficients for calculating heat propagation in the reservoir during heating of a shut-down well. If the reservoir heterogeneity, evaporation processes, and diffusion-capillary mass transfer of the saturating medium are not taken into account when describing thermophysical processes, significant discrepancies may be found between the observed data taken from the wells and the calculated values. Noticeable discrepancies were noted between the calculated temperatures at the well bottom and the values recorded at the fields. The observed differences in the well flow rate and the formation cooling period after treatment between the calculated and actual values generally turn out to be greater than the calculated ones. Consequently, the influence of phase transitions and diffusion-capillary effects on heat propagation throughout the heated reservoir is significant in a number of cases.

A model of a homogeneous thermally isotropic reservoir with infinite thickness and extension is proposed to which access is provided through a borehole of zero diameter. We assume that a heater with zero diameter and finite length h is placed in the borehole. Let a single-component liquid in thermodynamic equilibrium with the vapor contained in it and a gas insoluble in the liquid be used to fill the reservoir to saturation. The borehole is considered to be stopped when its initial temperature (T_0), pressure (p_0), and liquid saturation (θ_0) are uniformly distributed.

On the wall of the heater at the initial moment of time $t = 0$, the specific heat flow N is abruptly formed and its value maintained at constant. Due to the formation of such a stationary flow, the temperature T increases and the liquid begins to evaporate. The increase in temperature and intensification of evaporation in turn lead to an increase in the partial and total pressure p and a decrease in the saturation of the liquid phase θ.

When formulating the model, the following assumptions are also taken into account. The reservoir has a high degree of saturation. Since diffusion-capillary mass transfer of liquid and vapor prevails significantly over convective transfer, the latter can be neglected. The coefficients of diffusion-capillary mass transfer can be considered constant during the entire heating time and uniformly distributed everywhere in the considered region of heat flow formation. We will also assume that heat losses above and below the heater installation interval can be neglected due to their smallness compared to the power of the supported heat flow. The length of

the heater is considered to be so large that its dimensions affect the distribution of heat in the middle part of the heating interval. The formulated assumptions imply the heating of a fine-pored collector. The considered thermophysical characteristics in the middle part of the heating interval can also be assumed to be radially and axially symmetrically distributed in the plane. This reduces the mathematical formulation of the problem to a one-dimensional problem in which it is necessary to determine their nonstationary radial distributions.

This model allows for more accurate prediction of heat distribution in the reservoir, including phase transitions and diffusion-capillary effects. As a result, it is possible to more accurately estimate the temperature field, well flow rate after treatment, and duration of formation cooling.

As a result of such assumptions, the thermophysical characteristics of the formation can be considered as dependent on the radius r and time t: $T = T(r, t)$, $p = p(r, t)$, and $\theta = \theta(r, t)$. Their values at the initial moment of time are considered constant: $T(r, 0) = T_0$, $p(r, 0) = p_0$, and $\theta(r, 0) = \theta_0$.

For the convenience of the mathematical formulation of the model, we introduce such dimensionless parameters as dimensionless temperature:

$$T^* = \frac{T - T_0}{T};$$

dimensionless saturation of the liquid phase:

$$\theta^* = \frac{\theta_0 - \theta}{\theta_0};$$

dimensionless pressure resulting from the evaporation of the saturating liquid:

$$p^* = \frac{p - p_0}{p_0};$$

dimensionless radius:

$$R = \frac{r}{h};$$

dimensionless time (Fourier number):

$$Fo = \frac{\chi t}{h^2},$$

where χ is the thermal diffusivity (considered a constant value):

$$\chi = \frac{\lambda}{c\rho_e}.$$

Taking into account the above assumptions, the system of heat and mass transfer equations (1), (4), and (7) can be written in dimensionless form

$$\frac{\partial T^*}{\partial Fo} = \frac{\partial^2 T^*}{\partial R^2} + \frac{1}{R}\frac{\partial T^*}{\partial R} + \varepsilon\Pi_1\frac{\partial \theta^*}{\partial R}, \quad (8)$$

$$\frac{1}{Lu}\frac{\partial \theta^*}{\partial Fo} = \frac{\partial^2 \theta^*}{\partial R^2} + \frac{1}{R}\frac{\partial \theta^*}{\partial R} + \Pi_2\left(\frac{\partial^2 T^*}{\partial R^2} + \frac{1}{R}\frac{\partial T^*}{\partial R}\right), \quad (9)$$

$$\frac{\partial p^*}{\partial Fo} = -\frac{\varepsilon}{\Pi_3}\frac{\partial \theta^*}{\partial Fo}, \quad (10)$$

where $Fo > 0$ and $R > 0$ are the dimensionless time and radius, respectively; and Π_1, Π_2, Π_3, and Lu are the dimensionless coefficients described below. These quantities characterize the effect of mass transfer on the distribution of heat flows and vice versa, i.e., the effect of temperature on the saturation distribution. In particular, the Lykov criterion, defined as

$$Lu = \frac{a}{\chi\rho_e}$$

characterizes the intensity of diffusion-droplet mass transfer relative to diffusion heat transfer. If $Lu > 1$, then the thermal field due to diffusion-droplet mass transfer spreads faster than due to molecular thermal conductivity. In water-saturated sand and clay environments, its values are in the range of 0–3; in oil-saturated sands, $Lu \approx 1$.

The coefficient Π_1 defined as

$$\Pi_1 = \frac{mq\rho\Delta\theta}{c\Delta T}$$

characterizes the relationship between the amounts of heat spent on evaporation of the saturating liquid and on heating the formation (ΔT and $\Delta\theta$ are small changes in temperature and saturation in the volume element of the formation, respectively). In water-saturated sandy and clayey environments, its values are in the range 0.3–12.

The coefficient Π_2 defined as

$$\Pi_2 = \delta\frac{\Delta T}{\Delta\theta}$$

characterizes the increase in saturation $\Delta\theta$ due to a change in temperature by an amount ΔT. In water-saturated sandy and clayey environments, its values are in the range 0.1–0.9.

The coefficient Π_3 defined as

$$\Pi_3 = \beta\frac{\rho_e\Delta p}{\rho\Delta\theta},$$

characterizes the increase in saturation $\Delta\theta$ due to a change in pressure by an amount Δp. In water-saturated sandy and clayey environments, its values are in the range 0.2–0.7.

Equations (8)–(10) are supplemented by the initial conditions:

$$T^*\big|_{Fo=0} = 0,\ \theta^*\big|_{Fo=0} = 0,\ p^*\big|_{Fo=0} = 0, \quad (11)$$

as well as boundary conditions on the well axis:

$$\left.\frac{\partial T^*}{\partial R}\right|_{R=0} = -\frac{N}{2\pi\lambda}, \quad \left.\frac{\partial \theta^*}{\partial R}\right|_{R=0} = 0, \qquad (12)$$

and at infinity:

$$T^*\big|_{R\to\infty} = 0, \quad \theta^*\big|_{R\to\infty} = 0, \quad p^*\big|_{R\to\infty} = 0. \qquad (13)$$

Here, since the heat flux on the heater axis is considered constant, N/λ = const (12); moreover, the condition for the derivative of saturation corresponds to a constantly preserved maximum value of saturation being maintained on the well axis.

Thus, initial boundary value problem (8)–(13) is a mathematical formulation of the model of thermophysical processes in an oil reservoir during heating of a shut-in well. In order to solve this problem, a numerical algorithm and software package implementing it were developed.

3. DEVELOPMENT OF A COMPUTER IMPLEMENTATION OF THE MODEL

For the computer implementation of the model, an algorithm has been developed for solving the initial–boundary value problem for the system of equations (8)–(10) under boundary conditions (11)–(13) using finite-difference approximations of partial derivatives. Both explicit and implicit schemes with stability condition analysis were implemented.

The computer implementation of the model was carried out in the Python programming language using the NumPy (for working with data arrays and solving systems of equations) and Dash (for visualization of calculation results). The developed application consists of the following main components: data input module, equation system solution module, output and solution visualization module.

The module for solving a system of differential equations describing the thermal effect on a reservoir is implemented using finite-difference schemes.

The application also contains a data processing subsystem required to transform the data obtained from the equation solving module. First, explicit and implicit objects are generated to represent the equation solving strategies. Then, a solution method is called for each object using an appropriate technique to solve the system of equations. The program has the ability to output graphs displaying the results of calculations using explicit and implicit schemes. Each graph shows curves of temperature, saturation, and pressure. By switching between studying the curves and comparing data for the schemes, the user is able to better understand the processes occurring during thermal stimulation of the reservoir.

The application controls allow the user to view and analyze the results of solving equations using explicit and implicit methods in real time. Such controls allow setting several constants, including steps and integration limits for time and space variables. Using a drop-down list, the user can select output functions (temperature, saturation, and pressure), while the values of the control parameters are adjusted via interactive sliders. A specific value for another variable can be selected along with the variable (temporal or spatial) to build a graph of the distribution of thermophysical characteristics. The program provides the user with flexibility in setting up parameters by placing a panel with control elements on the left side of the main page and concentrating all settings in one screen location.

The set of sliders and drop-down menus is a component of the settings panel. A variable (temporal or geographic) can optionally be specified to serve as the basis for the abscissa axis of the graphs. Sliders may be used to change the numerical values of the 3 control parameters that affect the process. Another slider allows one to select the value of the spatial or temporal variable for visualization. The application interface tool is used to study the behavior of the distribution (temperature, saturation, or pressure) based on time or spatial variables. To achieve this goal, a graph can be produced with the values of the function under study plotted on the ordinate axis and the time or spatial variable plotted on the abscissa axis. In this case, the slider in the settings panel can be used to fix a certain value of the other variable. The resulting graphs display changes in the pressure, saturation, and temperature functions as a function of the spatial variable when the time variable is fixed. This focuses attention on how one variable, such as time or spatial coordinate, affects fixed parameters that are determined by another.

4. RESULTS AND DISCUSSION

As a result of numerical modeling using the developed software package, radial distribution curves were obtained (Fig. 1) along with time dependencies (Fig. 2) of temperature and saturation (by virtue of Eq. (10). A separate analysis of pressure does not make sense, since the corresponding dependencies are similar to saturation).

The dependencies of these quantities on the radius (Fig. 1) can be used to identify features of the spatial distribution of thermophysical characteristics. In particular, with increasing distance from the center, the temperature, saturation, and pressure decrease with a gradual slowdown. However, these radial dependencies of temperature and saturation are of different natures.

The temperature drops fairly quickly at small radii. Then its decrease slows down and gradually stops, reaching an equilibrium value already at relatively small values of the ratio of the radius to the length of the heater. Saturation (and pressure, by analogy) at small radii begins to decrease slowly, then quickly fall, and finally, at large distances, it reaches an equilibrium value and stops changing. This difference is due to the fact that the heat flow on the heater axis is maintained constant, and there is no saturation gradient on this axis.

The time dependencies of these quantities (Fig. 2) are used to identify the kinetic features of thermophysical processes. In particular, over time, temperature, saturation, and pressure increase with a gradual slowdown. However, at different radii, these dependencies are of a different nature. At small radii and short times, a sharp increase in temperature, saturation, and pressure is observed, followed by a rapid decline in the growth rates of these quantities. At large radius values, temperature, saturation, and pressure at first remain practically unchanged and then begin to grow slowly, gradually gaining growth speed. Such patterns correspond to upward convex sections of kinetic curves (at $R = 1$) and concave sections (at $R = 2$ and 3) on Fig. 2.

The parameters Π_1 and Π_2 are the most important in the model under consideration. They affect the saturation functions θ^* and temperature T. Varying these parameters implies actually going through different environments of the bed. Therefore, they are selected as the controlling thermophysical parameters of the model to evaluate the characteristics of processes in different soils.

If we accept $\Pi_1 = \Pi_2 = 0$, then this corresponds to the absence of mass transfer in the thermal physics problem. Vakhitov and Simkin [16] indicated that the mass transfer accounting leads to an increase in the radius of thermal influence, which, in turn, leads to higher calculated values of the formation cooling duration and an increase in flow rate after treatments. In accordance with the condition of maintaining thermal balance, an increase in the radius of thermal influence leads to a decrease in temperature on the wall and in the immediate vicinity of the well.

Heat transfer through a porous medium is controlled by the parameter Π_1, which reflects the effect of the saturation gradient on temperature change. This parameter is involved in Eq. (8), which describes the temperature change. As indicated by a higher value of Π_1, an increase in the rate of temperature change can be the result of an increase in the heat flow passing through the porous medium. The effect of the saturation gradient on temperature change is stronger the higher the value of Π_1.

The value of saturation is more significantly affected by the parameter Π_2. In addition, Π_2 affects temperature changes. The coefficient Π_2 associated with the influence of the temperature gradient on the change in saturation is included in the equation for saturation (9). The influence

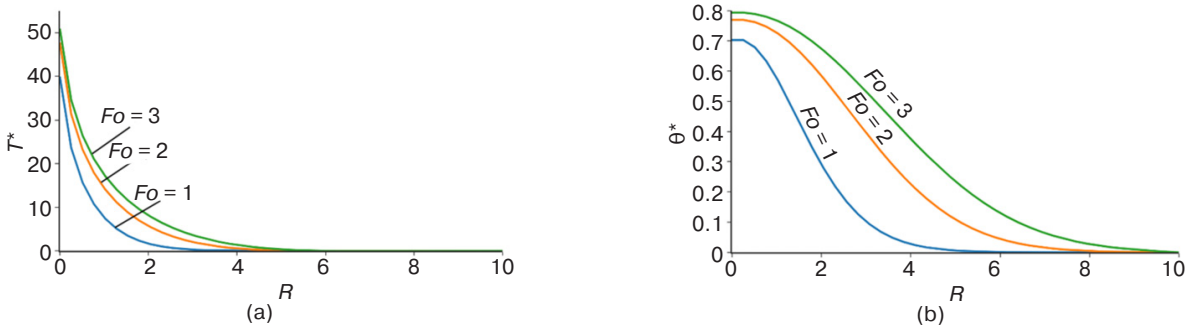


Fig. 1. Radial distributions of (a) temperature and (b) saturation at $\Pi_1 = 10$, $\Pi_2 = 0.1$, $\Pi_3 = 0.5$, Lu = 1, e = 0.28, l = 2.49 W/(m·K), and $N = 1000$ W/m

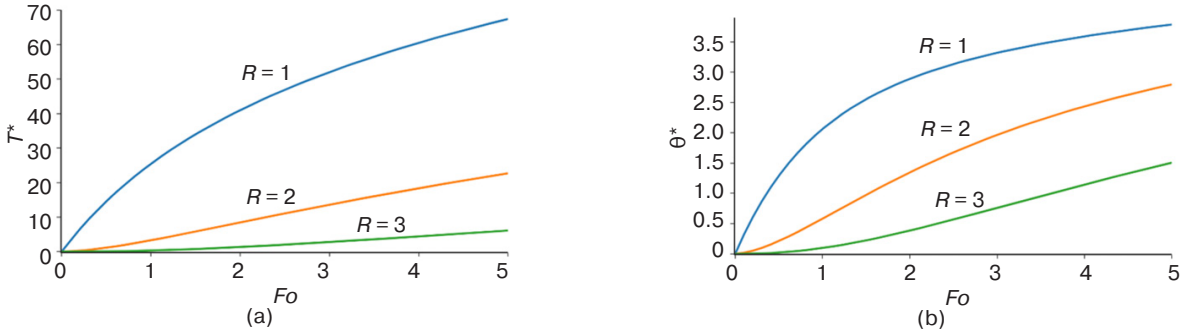


Fig. 2. Curves of (a) temperature kinetics and (b) saturation at parameter values as in Fig. 1

of the temperature gradient on the change in saturation is greater the higher the value of Π_2.

Thus, in the model, the parameters Π_1 and Π_2 determine the degree of relationship between temperature T^* and saturation θ^*. When modeling cooling after switching off the heater, these parameters can be used to adjust the degree of relationship between these two variables.

The results of the conducted series of numerical experiments were used to identify patterns of influence of control parameters on spatial distributions and kinetics of thermophysical characteristics. The influence of parameter Π_1 on temperature and saturation is illustrated in Figs. 3 and 4. An increase in the value of Π_1 leads to a more noticeable change in the saturation distribution along the radial coordinate than in temperature. From the point of view of the model structure, this is due to the fact that Π_1 is part of the coefficient in Eq. (1) for the derivative of saturation θ^* along the radial coordinate *R*. The radial distribution of saturation shifts more smoothly at large distances from the well and more steeply near it (small values of the radial coordinate *R*) as Π_1 increases. The saturation gradient has a greater effect on temperature fluctuations with an increase in Π_1, which can lead to a more rapid change in temperature in areas with a strong saturation gradient. From the physical point of view, sorting through the values of Π_1 corresponds to that of various fields in which the density and porosity of the soil and oil differ.

The simulation results shown in Fig. 3 show that a thousand-fold increase in the value of the parameter Π_1 leads to a very slight decrease in temperature (Fig. 3a) and a slight increase in saturation (Fig. 3b), which depends on the distance from the heater axis at a fixed point in time. Consequently, the obtained results indicate that the porosity and density of the soils and the liquid phase of the formation have little effect on the spatioradial distribution of temperature, saturation, and pressure in the formation.

The kinetic curves obtained as a result of modeling (Fig. 4) show that a thousand-fold increase in the value of the parameter Π_1 leads to an insignificant decrease in temperature (Fig. 4a) and an insignificant increase in saturation (Fig. 4b) at a fixed distance from the heater axis. Since this difference increases with increasing time, the influence of the porosity and density of soils and the liquid phase of the formation on the kinetics of temperature, saturation and pressure in the formation also begin to be noticeable at long time durations.

Physically, the enumeration of the values of Π_2 corresponds to the enumeration of various values of

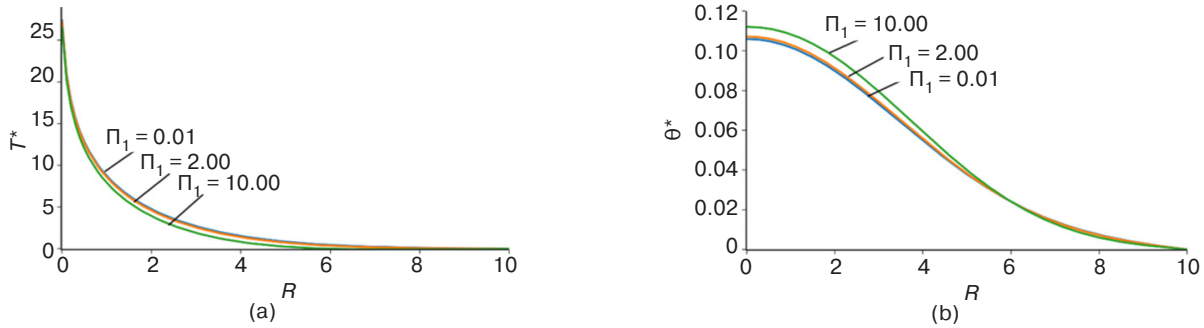


Fig. 3. Radial distributions of (a) temperature and (b) saturation at different values of the parameter Π_1 and fixed *Fo* = 2 (the other parameter values are the same as in Fig. 1)

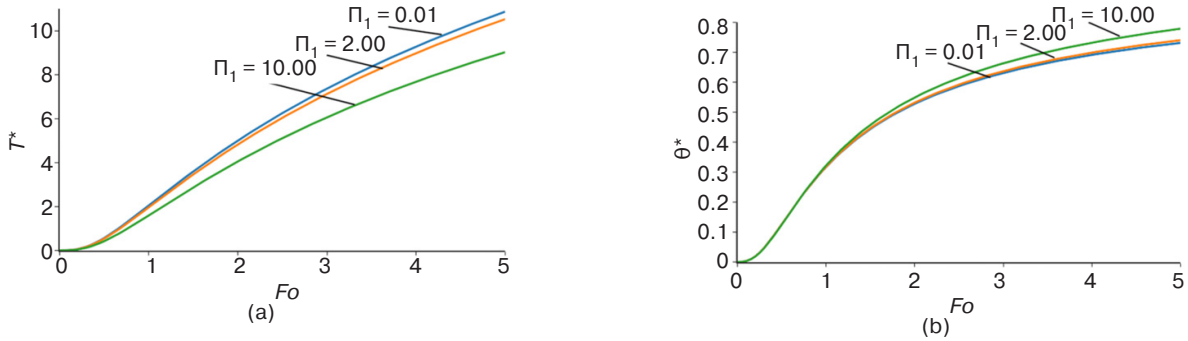


Fig. 4. Curves of (a) temperature kinetics and (b) saturation at different values of the parameter Π_1 and fixed *R* = 2 (the other parameter values are the same as in Fig. 1)

the thermogradient coefficient δ, which characterizes the transfer (flow) of moisture in the liquid phase of the oil-bearing formation. The modeling results (Fig. 5) show that a thousand-fold increase in the value of the parameter Π_2 leads to a very insignificant decrease in temperature (Fig. 5a), but to a significant increase in saturation and pressure (Fig. 5b) depending on the distance from the heater axis at a fixed point in time. An increase in the value of Π_2 leads to more rapid changes in the radial distribution of temperature T^* and saturation θ^*. The distributions of T^* and θ^* shift more smoothly at large distances from the well and more sharply near it as Π_2 increases. While saturation changes slightly at small Π_2, the saturation of the liquid phase of the oil formation begins to depend strongly on the distance from the heater axis at values of the thermogradient coefficient corresponding to water-saturated sands and clays.

Figure 6 illustrates the influence of the thermogradient coefficient, which can be used to evaluate the effect of moisture transfer in the liquid phase of the formation under the action of a temperature gradient in non-isothermal conditions on the kinetics of temperature and saturation. The growth of the thermogradient coefficient in the interval corresponding to water-saturated sands and clays leads to a noticeable decrease in temperature and an increase in the saturation of the liquid phase of the oil formation in such soils.

The influence of the elasticity coefficient of the vapor–gas mixture in the reservoir and the ratio of the skeleton and liquid phase densities are both characterized by the variation of the value of the parameter Π_3. This influence is particularly significant with regard to pressure (Fig. 7). The simulation results show that an increase in the value of the parameter Π_3 leads to a decrease in pressure in general. Although this decrease is almost imperceptible at short times, it becomes very noticeable at long times given a fixed distance from the heater axis (Fig. 7a). The same significant decrease with an increase in the elasticity coefficient of the vapor–gas mixture is observed at short distances from the heater axis at a fixed point in time (Fig. 7b). At large distances from the heater axis, a change in the parameter Π_3 ceases to affect the pressure in the reservoir.

The influence of the intensity of diffusion-capillary mass transfer relative to the diffusion heat transfer in the reservoir is characterized by the variation of the Lykov criterion *Lu*. Figure 8 shows the dependencies of the kinetics and radial distribution of saturation with variation of this criterion. The modeling results show that an increase in the value of the Lykov criterion leads to a sharp increase in saturation even at short

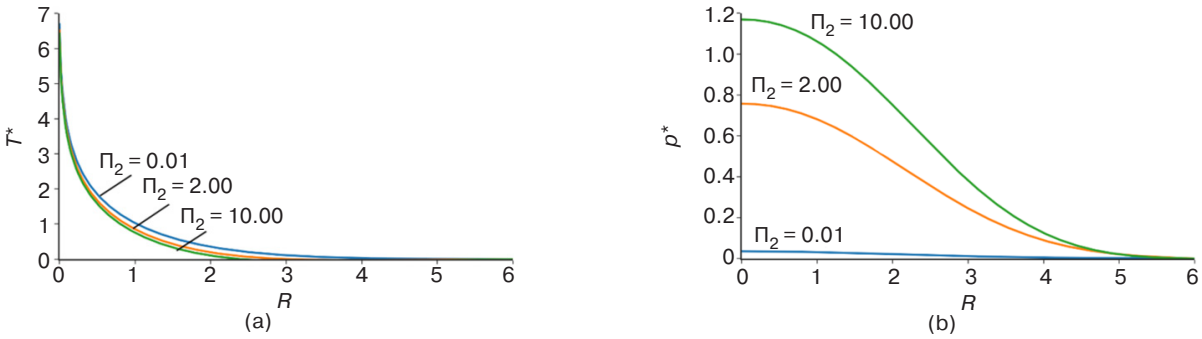


Fig. 5. Radial distributions of (a) temperature and (b) saturation at different values of the parameter Π_2 and fixed *Fo* = 2 (the other parameter values are the same as in Fig. 1)

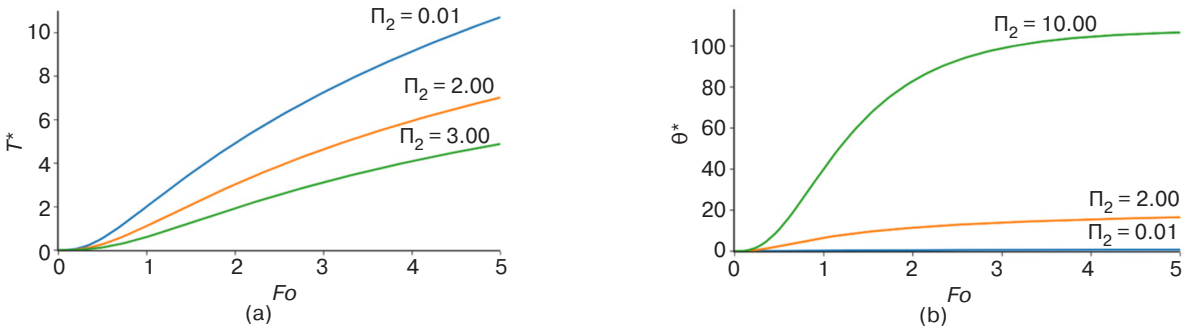


Fig. 6. Curves of (a) temperature kinetics and (b) saturation at different values of the parameter Π_2 and fixed *R* = 2 (the other parameter values are the same as in Fig. 1)

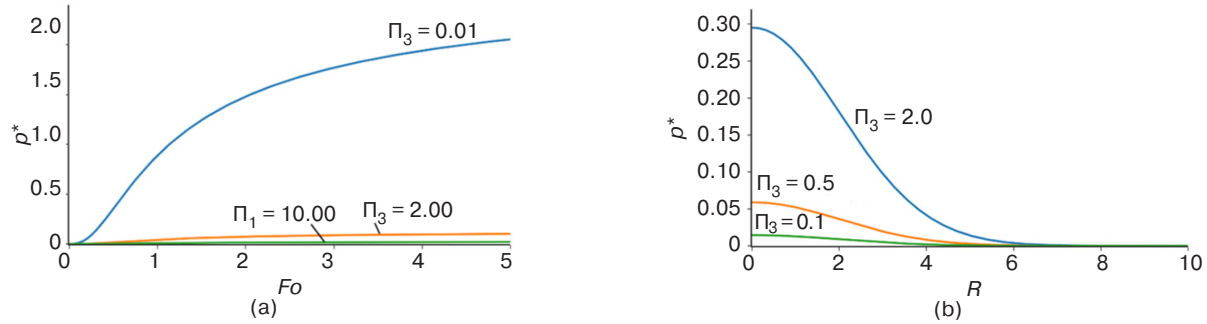


Fig. 7. Curves of (a) kinetics and (b) radial pressure distribution at different values of parameter Π_3 and fixed (a) $R = 2$ and (b) $Fo = 2$ (other parameter values as in Fig. 1)

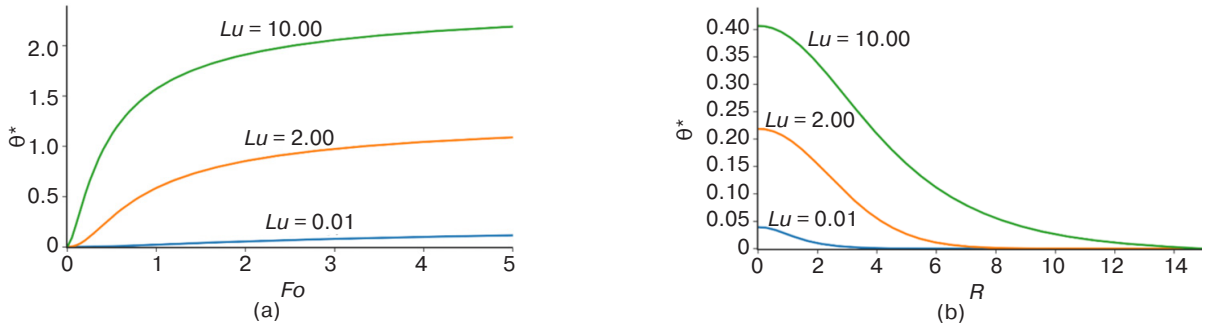


Fig. 8. Curves of (a) kinetics and (b) radial distribution of saturation at different values of the parameter *Lu* and fixed $R = 2$ (a) and (b) $Fo = 2$ (the remaining parameter values are the same as in Fig. 1)

times (Fig. 8a). An increase in the intensity of diffusion-capillary mass transfer relative to the diffusion heat transfer in the reservoir leads to an increase in saturation at a fixed point in time at short distances from the heater axis (Fig. 8b). At large distances from the heater axis, a change in the intensity under consideration does not affect saturation.

It should be noted that the variation of the Lykov criterion has virtually no effect on the temperature profiles. A slight decrease by fractions of a unit (in dimensionless temperature units T^*) was observed only at fairly long times (about 5 units in dimensionless time units of *Fo*).

Such a model parameter as the heat flow on the heater axis, as determined according to Eq. (12) by the ratio N/λ, also affects the value of the thermophysical characteristics but only as a numerical multiplier, i.e., temperature, saturation, and pressure are directly proportional to the specific heat flow on the heater axis.

CONCLUSIONS

The paper examines a mathematical model that describes the patterns of thermophysical processes in a formation during heating while taking into account the diffusion-capillary mass transfer of liquid and steam. The model, which provides a description of the occurring physical phenomena, takes into account the complex interaction of heat transfer processes and changes in the saturation of the liquid phase and pressure.

Software for the computer implementation of the model was developed using Python libraries (Pandas, Dash, NumPy) to provide convenient visualization and analysis of the modeling results. The inclusion of various modules for the numerical solution of the thermophysical initial–boundary value problem allows the selection of solution algorithms according to explicit or implicit schemes, along with discretization steps and other parameters for optimizing the computational procedure. The software contains a convenient user interface for entering parameters, displaying results in the form of tables and graphs, and exporting data.

Using the developed software, numerical experiments were conducted to study how various factors, such as reservoir properties, heater characteristics, and initial and boundary conditions, affect the reservoir cooling process. The analysis and interpretation of results showed that the spatial temperature distribution profile decreases quite quickly at small radii, followed by a slowdown in the decrease to an equilibrium value. The spatial saturation and pressure distribution profiles decrease at small radii and reach equilibrium values at large distances. Over time, the considered thermophysical characteristics monotonically increase.

It is shown that the porosity and density of soils and the liquid phase of the reservoir have a weak effect on the spatial distributions of temperature, saturation, and pressure in the reservoir. The influence of the porosity and density of soils and the liquid phase of the reservoir on the kinetics of these thermophysical characteristics is clearly noticeable at long time durations.

The results of the modeling showed that the saturation changes slightly at small values of the thermogradient coefficient; however, at values corresponding to water-saturated sands and clays, the saturation of the liquid phase of the oil reservoir begins to depend strongly on the distance from the heater axis. The growth of this coefficient in this range leads to a noticeable decrease in temperature and an increase in the saturation of the liquid phase of the oil reservoir in such soils.

It is confirmed that an increase in the coefficient of elasticity of the steam–gas mixture leads to a decrease in pressure at short distances from the heater axis, but that at long distances, the pressure stops changing. An increase in the intensity of diffusion-capillary mass transfer relative to the diffusion transfer of heat in the reservoir leads to a sharp increase in saturation at short distances from the heater axis; however, at longer distances, the saturation stops changing.

The results of this work may be useful for the oil industry when it is necessary to understand the patterns of thermal physical processes in the formation during heating to create more effective methods of managing the production process as a means of increasing its profitability. The proposed model and its analysis also expand the understanding of the possibilities of using mathematical and computer modeling of processes in the oil industry.

Authors' contributions

S.E. Savotchenko—problem statement, formulating the model, deriving equations, analyzing the results, writing the text of the article.

V.A. Zakharov—discretization of the model, developing the algorithm for a numerical method, developing the software for the model, computer experiments, and visualization.

REFERENCES

1. Gilmanov A.Ya., Kovalchuk T.N., Skoblikov R.M., Fedorov A.O., Khodzhiev Ye.N., Shevelev A.P. Analysis of the influence of thermophysical parameters of the reservoir and fluid on the process of cyclic steam stimulation. *Vestnik Tyumenskogo gosudarstvennogo universiteta. Seriya: Fiziko-matematicheskoe modelirovanie. Neft', gaz, energetika = Tyumen State University Herald. Physical and Mathematical Modeling. Oil, Gas, Energy*. 2023;9(3–35):6–27 (in Russ.). https://doi.org/10.21684/2411-7978-2023-9-3-6-27
2. Zagrivnyi E.A., Rudakov V.V., Bataev S.N. Electro thermal complex for thermal methods of extraction of high-viscosity oil. *Zapiski Gornogo instituta = Journal of Mining Institute*. 2004;158:226–229 (in Russ.).
3. Aminev D.A., Kravchenko M.N. Methods of thermal impact on the reservoir at the stage of well heating with pressure reduction. In: *SNK-2020: Proceedings of the Anniversary 70th Open International Student Scientific Conference of the Moscow Polytech*. Moscow. 2020. P. 198–201 (in Russ.).
4. Dumanskii Yu.G., Khlebnikov P.A., Mokropulo I.P., Gilaev G.G. Production of highly viscous oils by the method of thermal impact on reservoirs at the Okha field. In: *Development of Resources of Hard-to-Recover and High-Viscosity Oils: Conference Proceedings*. 1997. P. 142–165 (in Russ.).
5. Lapatin V.V. Analysis of the efficiency of thermal impact on a reservoir with highly viscous oil using a pair of simultaneously operating horizontal wells. *Aktual'nye problemy sotsial'no-gumanitarnogo i nauchno-tekhnicheskogo znaniya = Actual Problems of Social, Humanitarian and Scientific-Technical Knowledge*. 2023;4(35):18–20 (in Russ.).
6. Usmanov A.R. Thermal methods of influence on the formation. *Akademicheskii zhurnal Zapadnoi Sibiri = Academic Journal of West Siberia*. 2018;14(6–77):141–142 (in Russ.).
7. Gizzatullina A.A. Thermal impact on highly viscous oil in the reservoir using a pair of horizontal wells operating simultaneously. In: *12th All-Russian Congress on Fundamental Problems of Theoretical and Applied Mechanics: Conference Proceedings in 4 v*. V. 2. 2019. P. 1178–1179 (in Russ.). https://www.elibrary.ru/qfwigl
8. Salavatov T.Sh., Mamedov F.F. The prospects of using devices of alternative energy sources towards thermal treatment of formation and bottomhole zone. *Azerbaidzhanskoe neftyanoe khozyaistvo = Azerbaijan Oil Industry*. 2019;4:37–44 (in Russ.).
9. Yangirov R.R. Dynamics of the temperature field in the reservoir during thermal impact on the productive reservoir on the example of the Russian field. In: West Siberian Oil and Gas Congress: *Collection of Scientific Papers of the 13 International Scientific and Technical Congress of the Student Section of the Society of Petroleum Engineers* (*SPE*). 2019. P. 181–183 (in Russ.). https://www.elibrary.ru/zgqxbb

10. Igtisanova G.R., Gabdullina G.I. Calculation of the parameters of the kust pump station with autonomous energy provision and heat, water-gas reservoir stimulation. In: *Actual Issues of Higher Education – 2020: Proceedings of the All-Russian Scientific and Methodological Conference* (*with International Participation*). 2020. P. 230–237 (in Russ.). https://www.elibrary.ru/dfeidr
11. Gizzatullina A.A., Fatkullin A.A., Abdulmanov A.A. Mathematical modeling of heat influence on a high-viscous oil in a layer through a horizontal well. In: *Contemporary technologies in the oil and gas industry – 2020: Collection of Works of the International Scientific and Technical Conference.* 2020. P. 45–49 (in Russ.). https://www.elibrary.ru/fvllkz
12. Kartashov E.M., Kudinov V.A. *Analiticheskaya teoriya teploprovodnosti i prikladnoi termouprugosti* (*Analytical Theory of Heat Conduction and Applied Thermoelasticity*). Moscow: URSS; 2018. 656 p. (in Russ.).
13. Kartashov E.M. Model representations of heat shock in terms of dynamic thermal elasticity. *Russ. Technol. J*. 2020;8(2): 85–108 (in Russ.). https://doi.org/10.32362/2500-316X-2020-8-2-85-108
14. Sayakhov F.L., Fatykhov M.A., Dyblenko V.M., Simkin E.M. Calculation of the main indicators of the process of high-frequency heating of the bottomhole zone of oil wells. *Izvestiya vysshikh uchebnykh zavedenii. Seriya Neft' i gaz = Oil and Gas Studies*.1977;6:15–21 (in Russ.). https://elibrary.ru/nzemgp
15. Surguchev M.L., Simkin E.M., Zhdanov S.A. Influence of methods of thermophysical impact on bottomhole zones on increasing oil recovery. *Neftyanoe khozyaistvo = Oil Industry*. 1977;10:48–50 (in Russ.).
16. Vakhitov G.G., Simkin E.M. *Ispol'zovanie fizicheskikh polei dlya izvlecheniya nefti iz plastov* (*The Use of Physical Fields to Extract Oil from Reservoirs*). Moscow: Nedra; 1985. 231 p. (in Russ.).
17. Durkin S.M., Menshikova I.N. Assessment of the impact of method of accounting of non-reservoir rocks in the process of thermal stimulation by numerical simulation. *Inzhener-neftyanik*. 2017;3:38–41 (in Russ.). https://elibrary.ru/zhneuf
18. Dieva N.N., Kravchenko M.N., Nabiullina A.A. Justification based on numerical modeling of the choice of thermal impact methods on kerogen-containing reservoirs. In: Actual Problems of Oil and Gas Geology in Siberia: *Proceedings of the 2nd All-Russian Scientific Conference of Young Scientists and Students Dedicated to the 85th Anniversary of Academician A.E. Kontorovich*. 2019. P. 37–39 (in Russ.). https://elibrary.ru/mukjjh
19. Tupysev M.K. Dynamics of hydrate formation in near-wellbore zone of wells while the development of low-temperature gas deposits. *Georesursy, Geoenergetika, Geopolitika.* 2010;2(2):20 (in Russ.). https://elibrary.ru/sikwul
20. Sharafutdinov R.F., Kanafin I.V., Khabirov T.R. Numerical Investigation of the temperature field in a multiple-zone well during gas-cut oil motion. *J. Appl. Mech. Tech. Phys*. 2019;60(5):889–898. https://doi.org/10.1134/S0021894419050122
[Original Russian Test: Sharafutdinov R.F., Kanafin I.V., Khabirov T.R. Numerical Investigation of the temperature field in a multiple-zone well during gas-cut oil motion. *Prikladnaya mekhanika i tekhnicheskaya fizika*. 2019;60(5–357):125–135 (in Russ.). https://doi.org/10.15372/PMTF20190512]
21. Ramazanov A.Sh., Parshin A.V. Temperature distribution in oil-water-saturated reservoir with account of oil degassing. *Neftegazovoe delo = Oil and Gas Business.* 2006;1:22 (in Russ.). https://elibrary.ru/twwnpx
22. Ramazanov A.S., Parshin A.V. Analytical model of temperature variations during the filtration of gas-cut oil. *High Temp*. 2012;50(4):567–569. https://doi.org/10.1134/S0018151X12040189
[Original Russian Test: Ramazanov A.Sh., Parshin A.V. Analytical model of temperature variations during the filtration of gas-cut oil. *Teplofizika vysokikh temperatur*. 2012;50(4):606–608 (in Russ.). https://elibrary.ru/kovvto]
23. Shagapov V.Sh., Tazetdinov B.I. Modeling the dynamics of formation and decomposition of gas hydrate particles during their surfacing in water. *Vestnik Samarskogo gosudarstvennogo universiteta. Estestvennonauchnaya seriya = Vestnik of Samara University. Natural Science Series.* 2013;9-2(110):133–139 (in Russ.). https://elibrary.ru/rxwiiv
24. Ramazanova E.N., Aliverdiev A.A., Grigor'ev E.B., Zarichnyak Yu.P., Aliev R.M., Beibalaev V.D. Temperature field of an oil reservoir considering the effect of thermal methods of impact and thermal conductivity of rocks. *Vesti gazovoi nauki.* 2023;2(54):68–73 (in Russ.). https://elibrary.ru/wrwewp
25. Gil'manov A.Y., Shevelev A.P., Lagunov P.S., et al. Influence of the Thermophysical Properties of the Reservoir and Fluid on the Technological Parameters of the Cyclic Steam Stimulation. *J. Eng. Phys. Thermophy.* 2023;96(5):1311–1319. https://doi.org/10.1007/s10891-023-02797-8
[Original Russian Test: Gil'manov A.Ya., Shevelev A.P., Lagunov P.S., Gulyaev P.N., Petukhov A.S., Lyutoev P.A. Influence of the Thermophysical Properties of the Reservoir and Fluid on the Technological Parameters of the Cyclic Steam Stimulation. *Inzhenerno-fizicheskii zhurnal.* 2023;96(5):1323–1331 (in Russ.).]
26. Efimtsev S.V., Nustrov B.C., Okhezin S.P., Podoplelov V.V. Some problems of filtration in deformable media. *Izvestiya Ural'skogo gosudarstvennogo universiteta. Seriya Matematika i mekhanika.* 2003;5(26):66–76 (in Russ.).
27. Makarychev S.V., Mazirov M.A. *Teplofizika pochv: metody i svoistva* (*Thermal Physics of Soils: Methods and Properties*). Suzdal; 1996. V. 1. 231 p. (in Russ.).
28. Laptev A.G., Nikolaev N.A., Basharov M.M. *Metody intensifikatsii i modelirovaniya teplomassoobmennykh protsessov* (*Methods of Intensification and Modeling of Heat and Mass Transfer Processes*). Moscow: Teplotekhnik; 2011. 287 p. (in Russ.).
29. Goldobin D.S., Krauzin P.V. Saturation of aquifers with two-component gas mixture. *Vestnik PNIPU. Mekhanika = PNRPU Mechanics Bulletin*. 2013;4:33–41 (in Russ.).
30. Goldobin D.S., Krauzin P.V. Formation of bubbly horizon in liquid-saturated porous medium by surface temperature oscillation. *Phys. Rev. E*. 2015;92(6):063032(1–8). http://dx.doi.org/10.1103/PhysRevE.92.063032

СПИСОК ЛИТЕРАТУРЫ

1. Гильманов А.Я., Ковальчук Т.Н., Скобликов Р.М., Фёдоров А.О., Ходжиев Ё.Н., Шевелёв А.П. Анализ влияния теплофизических параметров пласта и флюида на процесс пароциклического воздействия на нефтяные пласты. *Вестник Тюменского государственного университета. Серия: Физико-математическое моделирование. Нефть, газ, энергетика*. 2023;9(3–35):6–27. https://doi.org/10.21684/2411-7978-2023-9-3-6-27
2. Загривный Э.А., Рудаков В.В., Батаев С.Н. Электротермический комплекс для тепловых методов добычи высоковязкой нефти. *Записки Горного института.* 2004;158:226–229.
3. Аминев Д.А., Кравченко М.Н. Методы теплового воздействия на пласт на этапе прогрева скважины с понижением давления. В сб.: *СНК-2020: материалы Юбилейной LXX открытой международной студенческой научной конференции Московского Политеха*. Москва. 2020. С. 198–201.
4. Думанский Ю.Г., Хлебников П.А., Мокропуло И.П., Гилаев Г.Г. Добыча высоковязких нефтей методом теплового воздействия на пласты на месторождении Оха. В сб.: *Освоение ресурсов трудноизвлекаемых и высоковязких нефтей: сборник трудов конференции*. 1997. С. 142–165.
5. Лапатин В.В. Анализ эффективности теплового воздействия на пласт с высоковязкой нефтью с помощью пары одновременно работающих горизонтальных скважин. *Актуальные проблемы социально-гуманитарного и научно-технического знания*. 2023;4(35):18–20.
6. Усманов А.Р. Тепловые методы воздействия на пласт. *Академический журнал Западной Сибири.* 2018;14(6–77):141–142.
7. Гиззатуллина А.А. Тепловое воздействие на высоковязкую нефть в пласте с помощью пары горизонтальных скважин работающих одновременно. В сб.: *XII Всероссийский съезд по фундаментальным проблемам теоретической и прикладной механики: Сборник трудов конференции в 4 томах.* Т. 2. 2019. С. 1178–1179. https://www.elibrary.ru/qfwigl
8. Салаватов Т.Ш., Мамедов Ф.Ф. Перспективы использования установок альтернативных источников энергии с целью теплового воздействия на пласт и призабойную зону. *Азербайджанское нефтяное хозяйство*. 2019;4:37–44.
9. Янгиров Р.Р. Динамика температурного поля в пласте при тепловом воздействии на продуктивный пласт на примере русского месторождения. В сб.: *Западно-Сибирский нефтегазовый конгресс*: *Сборник научных трудов XIII Международного научно-технического конгресса студенческого отделения общества инженеров-нефтяников – Society of Petroleum Engineers* (*SPE*). 2019. С. 181–183. https://www.elibrary.ru/zgqxbb
10. Игтисамова Г.Р., Габдуллина Г.И. Расчет параметров кустовой насосной станции с автономным энерго-обеспечением и тепловым, водогазовым воздействием на пласт. В сб.: *Актуальные вопросы высшего образования – 2020: Материалы Всероссийской научно-методической конференции* (*с международным участием*). 2020. С. 230–237. https://www.elibrary.ru/dfeidr
11. Гиззатуллина А.А., Фаткуллин А.А., Абдульманов А.А. Математическое моделирование теплового воздействия на высоковязкую нефть в пласте через горизонтальную скважину. В сб.: *Современные технологии в нефтегазовом деле – 2020: сборник трудов международной научно-технической конференции*. 2020. С. 45–49. https://www.elibrary.ru/fvllkz
12. Карташов Э.М., Кудинов В.А. *Аналитическая теория теплопроводности и прикладной термоупругости*. М.: URSS; 2018. 656 с.
13. Карташов Э.М. Модельные представления теплового удара в динамической термоупругости. *Russ. Technol. J.* 2020;8(2):85–108. https://doi.org/10.32362/2500-316X-2020-8-2-85-108
14. Саяхов Ф.Л., Фатыхов М.А., Дыбленко В.М., Симкин Э.М. Расчет основных показателей процесса высокочастотного нагрева призабойной зоны нефтяных скважин. *Изв. вузов. Серия Нефть и газ.*1977;6:15–21. https://elibrary.ru/nzemgp
15. Сургучев М.Л., Симкин Э.М., Жданов С.А. Влияние методов теплофизического воздействия на призабойные зоны на увеличение нефтеотдачи пласта. *Нефтяное хозяйство*. 1977;10:48–50.
16. Вахитов Г.Г., Симкин Э.М. *Использование физических полей для извлечения нефти из пластов*. М.: Недра; 1985. 231 с.
17. Дуркин С.М., Меньшикова И.Н. Оценка влияния способа учета пород-неколлекторов на процесс теплового воздействия на пласт при численном моделировании. *Инженер-нефтяник*. 2017;3:38–41. https://elibrary.ru/zhneuf
18. Диева Н.Н., Кравченко М.Н., Набиуллина А.А. Обоснование на основе численного моделирования выбора методов теплового воздействия на керогеносодержащие пласты. В сб.: *Актуальные проблемы геологии нефти и газа Сибири*: *Материалы 2-й Всероссийской научной конференции молодых ученых и студентов, посвященной 85-летию академика А.Э. Конторовича.* 2019. С. 37–39. https://elibrary.ru/mukjjh
19. Тупысев М.К. Динамика гидратообразования в призабойной зоне скважин при разработке низкотемпературных газовых залежей. *Георесурсы, геоэнергетика, геополитика*. 2010;2(2):20. https://elibrary.ru/sikwul
20. Шарафутдинов Р.Ф., Канафин И.В., Хабиров Т.Р. Численное исследование температурного поля в скважине с многопластовой системой при движении газированной нефти. *Прикладная механика и техническая физика*. 2019;60(5–357): 125–135. https://doi.org/10.15372/PMTF20190512
21. Рамазанов А.Ш., Паршин А.В. Температурное поле в нефте-водонасыщенном пласте с учетом разгазирования нефти. *Нефтегазовое дело*. 2006;1:22. https://elibrary.ru/twwnpx
22. Рамазанов А.Ш., Паршин А.В. Аналитическая модель температурных изменений при фильтрации газированной нефти. *Теплофизика высоких температур*. 2012;50(4):606–608. https://elibrary.ru/kovvto
23. Шагапов В.Ш., Тазетдинов Б.И. Моделирование динамики образования и разложения газогидратных частиц при их всплытии в воде. *Вестник Самарского государственного университета. Естественнонаучная серия*. 2013;9-2(110): 133–139. https://elibrary.ru/rxwiiv

24. Рамазанова Э.Н., Аливердиев А.А., Григорьев Е.Б., Заричняк Ю.П., Алиев Р.М., Бейбалаев В.Д. Температурное поле нефтяного паста с учетом влияния тепловых методов воздействия и теплопроводности горных пород. *Вести газовой науки*. 2023;2(54):68–73. https://elibrary.ru/wrwewp
25. Гильманов А.Я., Шевелёв А.П., Лагунов П.С., Гуляев П.Н., Петухов А.С., Лютоев П.А. Влияние теплофизических свойств пласта и флюида на технологические параметры пароциклического воздействия. *Инженерно-физический журнал*. 2023;96(5):1323.
26. Ефимцев С.В., Нустров В.С., Охезин С.П., Подоплелов В.В. Некоторые задачи фильтрации в деформируемых средах. *Известия УрГУ. Серия Математика и механика*.2003;5(26):66–76. https://elibrary.ru/vlomwf
27. Макарычев С.В., Мазиров М.А. *Теплофизика почв: методы и свойства*. Суздаль; 1996. Т. 1. 231 с.
28. Лаптев А.Г., Николаев Н.А., Башаров М.М. *Методы интенсификации и моделирования тепломассообменных процессов*. М.: Теплотехник; 2011. 287 с.
29. Голдобин Д.С., Краузин П.В. Насыщение затопленных почв двухкомпонентной смесью газов. *Вестник ПНИПУ. Механика*. 2013;4:33–41.
30. Goldobin D.S., Krauzin P.V. Formation of bubbly horizon in liquid-saturated porous medium by surface temperature oscillation. *Phys. Rev. E*. 2015;92(6):063032(1–8). http://dx.doi.org/10.1103/PhysRevE.92.063032

About the Authors

Sergey E. Savotchenko, Dr. Sci. (Phys.-Math.), Associate Professor, Professor, High Mathematics Department, Institute for Advanced Technologies and Industrial Programming, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia); Professor, Department of Higher Mathematics and Physics, Sergo Ordzhonikidze Russian State University for Geological Prospecting (MGRI) (23, Miklukho-Maklaya ul., Moscow, 117997 Russia). E-mail: savotchenkose@mail.ru. Scopus Author ID 6603577988, RSCI SPIN-code 2552-4344, https://orcid.org/0000-0002-7158-9145

Vasily A. Zakharov, Postgraduate Student, Department of Higher Mathematics and Physics, Sergo Ordzhonikidze Russian State University for Geological Prospecting (MGRI) (23, Miklukho-Maklaya ul., Moscow, 117997 Russia). E-mail: vasilyblack10@gmail.com. https://orcid.org/0009-0008-1978-2993

Об авторах

Савотченко Сергей Евгеньевич, д.ф.-м.н., доцент, профессор кафедры высшей математики, Институт перспективных технологий и индустриального программирования, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78); профессор кафедры высшей математики и физики, ФГБОУ ВО «Российский государственный геологоразведочный университет им. Серго Орджоникидзе» (МГРИ) (117997, Россия, Москва, ул. Миклухо-Маклая, д. 23). E-mail: savotchenkose@mail.ru. Scopus Author ID 6603577988, SPIN-код РИНЦ 2552-4344, https://orcid.org/0000-0002-7158-9145

Захаров Василий Александрович, аспирант, кафедра высшей математики и физики, ФГБОУ ВО «Российский государственный геологоразведочный университет им. Серго Орджоникидзе» (МГРИ) (117997, Россия, Москва, ул. Миклухо-Маклая, д. 23). E-mail: vasilyblack10@gmail.com. https://orcid.org/0009-0008-1978-2993

Translated from Russian into English by Vladislav Glyanchenko
Edited for English language and spelling by Thomas A. Beavitt

Economics of knowledge-intensive and high-tech enterprises and industries. Management in organizational systems

Экономика наукоемких и высокотехнологичных предприятий и производств. Управление в организационных системах

UDC 005.7+621.5+621.316.7
https://doi.org/10.32362/2500-316X-2025-13-5-119-132
EDN TQKWPM

RESEARCH ARTICLE

Evaluation of the project based on the theory of fuzzy sets and the concept of fuzzy logic. Method and methodology

Viktor V. Sidorin @

Institute of Testing and Certification of Weapons and Military Equipment, Moscow, 111524 Russia
@ *Corresponding author, e-mail: wwsid@yandex.ru*

• **Submitted:** 14.02.2025 • **Revised:** 14.04.2025 • **Accepted:** 24.07.2025

Abstract

Objectives. The work sets out to develop a method and methodology for evaluating project activities, including research and development work. The development of this method and its associated methodology is relevant due to the need to provide an analytical assessment of a project based on its main performance indicators, such as the uniqueness of the results and the resources required to produce them. This assessment must take into account both the requirements of the customer and the capabilities of the prospective performers to make an informed decision on its formulation. Currently, the best-known methods for evaluating a planned project are based on economic efficiency. However, the approaches taken by customers and contractors are often different and sometimes contradictory. For example, a customer may minimize the risk of failing to achieve the project goal by setting appropriate requirements and resource costs, while a contractor minimizes the same risks by increasing the requested time and material resources, as well as by adjusting the requirement criteria for the project results, which are based on their ability to fulfill them. The concept of fuzzy sets allows various assessment approaches to be combined to provide informed and coordinated decision-making regarding the feasibility and expediency of setting up and executing a project.

Methods. The developed method and methodology for project evaluation are based on the theory of fuzzy sets and the concept of fuzzy logic, using membership functions to model project parameter estimates.

Results. The project assessment method and implementation methodology were developed, taking into account customer requirements and the capabilities of the potential contractor, to enable rational decision-making regarding its formulation.

Conclusions. The method may be applied to establish the optimal requirements and conditions for project implementation in a balanced manner. In contrast to expert methods, the presented analytical method and methodology provide a higher objectivity of the generalized assessment, validity, and effectiveness in making and implementing managerial decisions. The methodology's universal criteria and procedures make it suitable for application to various kinds of assessment and in various areas of project activity, including research and development work.

Keywords: project, theory of fuzzy sets, linguistic variables, membership functions, fuzzification, defuzzification

For citation: Sidorin V.V. Evaluation of the project based on the theory of fuzzy sets and the concept of fuzzy logic. Method and methodology. *Russian Technological Journal.* 2025;13(5):119–132. https://doi.org/10.32362/2500-316X-2025-13-5-119-132, https://www.elibrary.ru/TQKWPM

Financial disclosure: The author has no financial or proprietary interest in any material or method mentioned.

The author declares no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Оценка проекта на основе теории нечетких множеств и концепции нечеткой логики. Метод и методика

В.В. Сидорин [@]

Институт испытаний и сертификации вооружения и военной техники, Москва, 111524 Россия
[@] *Автор для переписки, e-mail: wwsid@yandex.ru*

• **Поступила:** 14.02.2025 • **Доработана:** 14.04.2025 • **Принята к опубликованию:** 24.07.2025

Резюме

Цели. Цель работы – разработка метода и методики оценки проектной деятельности, включая научно-исследовательские и опытно-конструкторские работы. Актуальность разработки обусловлена потребностью в аналитической оценке проекта по его основным показателям (уникальности результатов и требуемым ресурсам), учитывающей как требования заказчика, так и возможности предполагаемых исполнителей для принятия согласованного решения по его постановке. В настоящее время основой известных методов оценки планируемого проекта является экономическая эффективность. При этом подходы к оценке проекта заказчика и исполнителя различны и противоречивы. Заказчик в своих требованиях к проекту и затратам ресурсов минимизирует риски недостижения цели проекта. Исполнитель те же риски минимизирует увеличением запрашиваемых временных и материальных ресурсов, а также своими требованиями к результатам проекта, в основе которых – его возможности их выполнить. Объединить различные подходы к оценке для обоснованного и согласованного принятия решений относительно возможности и целесообразности постановки и выполнения проекта позволяет концепция нечетких множеств.

Методы. В основе разработанных метода и методики оценки проектов – теория нечетких множеств и концепция нечеткой логики с моделированием оценок параметров проекта с помощью функций принадлежности.

Результаты. Разработаны метод и реализующая его методика оценки проекта, учитывающие требования заказчика и возможности предполагаемого исполнителя для обоснованного принятия решений в отношении его постановки.

Выводы. Применение метода позволяет сбалансированно установить оптимальные требования и условия для осуществления проекта. В отличие от экспертных методов представленные аналитический метод и методика обеспечивают более высокую объективность обобщенной оценки, обоснованность и эффективность в принятии и реализации управленческих решений. Универсальность критериев и процедур методики содержит в себе возможность ее применения в оценках различных видов и по различным направлениям проектной деятельности, включая научно-исследовательские и опытно-конструкторские работы.

Ключевые слова: проект, теория нечетких множеств, лингвистические переменные, функции принадлежности, фаззификация, дефаззификация

Для цитирования: Сидорин В.В. Оценка проекта на основе теории нечетких множеств и концепции нечеткой логики. Метод и методика. *Russian Technological Journal*. 2025;13(5):119–132. https://doi.org/10.32362/2500-316X-2025-13-5-119-132, https://www.elibrary.ru/TQKWPM

Прозрачность финансовой деятельности: Автор не имеет финансовой заинтересованности в представленных материалах или методах.

Автор заявляет об отсутствии конфликта интересов.

INTRODUCTION

A project differs from operational activities in three main ways: its goal is to produce a unique result that has never been created before; it has a limited time frame; and it has limited resources. Risk factors that may prevent the project from achieving its goal included a lack of prior information or experience of performing similar work or creating a similar product. It is important to note that these risks apply to both the project contractor and the customer [1–3]. The repercussions of a contractor's erroneous evaluation of their competencies are twofold: firstly, in failing to attain the stipulated objective, and secondly, in being unable to conclude the project. In order to mitigate risk, contractors often overestimate the resources and deadlines required for project work and underestimate the level of resources necessary for the development stage [4–6]. Customers, who are willing to expend their own resources, generally seek to achieve the required result with fewer resources and in a shorter timeframe than contractors would like. Thus, an objective evaluation of the project that takes into account both the customer's requirements and the contractor's capabilities could foster a more harmonious relationship between the two parties.

One possible resolution to this issue consists in an assessment method that incorporates assessments of the same object by both parties to resolve the conflict of interests between them. This possibility is realized through an approach based on fuzzy logic and fuzzy set theory. This approach is based on the use and processing of large amounts of information using digital information technology methods and tools to obtain a variety of solutions for different evaluation criteria.

The need to use fuzzy logic methodology is based on the requirement to achieve more reliable assessment outcomes by using analytical methods that are sensitive to changes in the project activity being evaluated. These include individual projects, programs and project portfolios, as well as research and development work and preliminary designs. The results of assessments employing such analytical methods to evaluate proposed project activities according to their three attributes (criteria and characteristic features) have the potential to serve as the basis for joint decisions on setting and implementing project activities by interacting parties (customers and contractors).

The reliability of assessments is pivotal in determining the adequacy of decisions taken and the effectiveness of project management. To achieve this, assessment results must demonstrate sensitivity to changes in project parameters that lead to deviations from the set objectives. This establishes the requirements for the qualitative and quantitative indicators used to assess the project. Here it is important to note that methods based on measured data or expert opinion do not always allow for such an assessment [7–11].

1. MODEL FOR FORECASTING PROJECT PARAMETERS BASED ON THEIR FUZZY MODELING FOR MANAGEMENT DECISION-MAKING

The purpose of the method is to determine the feasibility of project activities based on a forecast assessment of three parameters comprising the novelty and originality of the planned results, the adequacy of the allocated resources and deadlines, and the optimal use and planning of these resources. These assessments, which are common to all interacting parties (customers and contractors), are characterized by ambiguity, the possibility of multiple responses, and different perceptions of their significance. This leads to decisions being made that vary in their effectiveness. The concept of fuzzy sets enables different assessments to be combined for informed and coordinated decision-making regarding the feasibility and advisability of setting up and implementing a project. The presented method involves modelling the customer's and contractor's project assessments using membership functions within a specific fuzzy set, thereby generating a generalized assessment.

The result of applying this method should be an assessment of the project in terms of its compliance with the customer's requirements and the contractor's capabilities. This should be followed by a decision on the

feasibility of implementation that satisfies both parties. Focusing on using artificial intelligence technologies to manage the project at all stages—from establishing feasibility to selecting optimal scientific, technical, and organizational solutions—requires mathematical models to be developed for each stage of project activity [12–14].

Fuzzy set theory enables the structuring of project criteria, parameters, and characteristics that cannot be assessed using a single scale and that can be perceived differently by the two interacting parties: contractors and project customers. The need to evaluate these parameters, which do not have a single, unambiguous rating scale, can be met by using the provisions of fuzzy set theory to determine whether they meet the criteria of the category "Meets requirements and satisfies participants." This can be achieved by modelling the project evaluation as a fuzzy system that links the impact of project assessments in terms of novelty, originality, sufficiency of resources, and deadlines for work completion according to customer requirements and contractor capabilities. This modelling process involves three stages: first, creating a mental model of the project; second, creating a verbal model of the project; and third, developing a fuzzy linguistic model of the project. Input parameters for such a fuzzy model can be obtained in three ways: through expert assessment of the parameters of the modelled project; by constructing self-tuning fuzzy models based on system input and output measurements; and by constructing self-organizing and self-tuning fuzzy models based on system input and output measurements[1] [15, 16].

The mental model of a project is a vision of the activities involved in creating a new, unique product within a limited timeframe and with limited resources. The verbal model, which describes the project's main properties and distinctive features in clear terms, is then transformed into a fuzzy linguistic model to facilitate informational interaction between project participants.

The consistency of two factors should be ensured by the fuzzy linguistic system and its model: the customer's requirements and the contractor's capabilities. The assessment is based on a joint assessment model that uses fuzzy logic methods, in which the customer and the prospective contractor evaluate the predicted project results.

The customer and contractor comprising the two interacting parties are considered as a fuzzy system [17–19]. The inputs for analysis in the fuzzy model are the project parameters predicted by each of the interacting parties. These inputs include the degree of originality of the scientific, technical, or other planned results of the project, the level of adequacy of resource support for the project work, and the adequacy of time for the project work. The analysis provides a general assessment of the project's alignment with the customer's needs and the contractor's capabilities. This requires a solution to the problem of combining the two independent assessments from each party. Based on this assessment, a decision is made on whether to implement the project.

The structure and substance of the project predictive assessment method, which is based on fuzzy set theory, fuzzy logic, and fuzzy modelling of its characteristics, is used to make decisions on project initiation and management. The method comprises the following stages (Fig. 1):

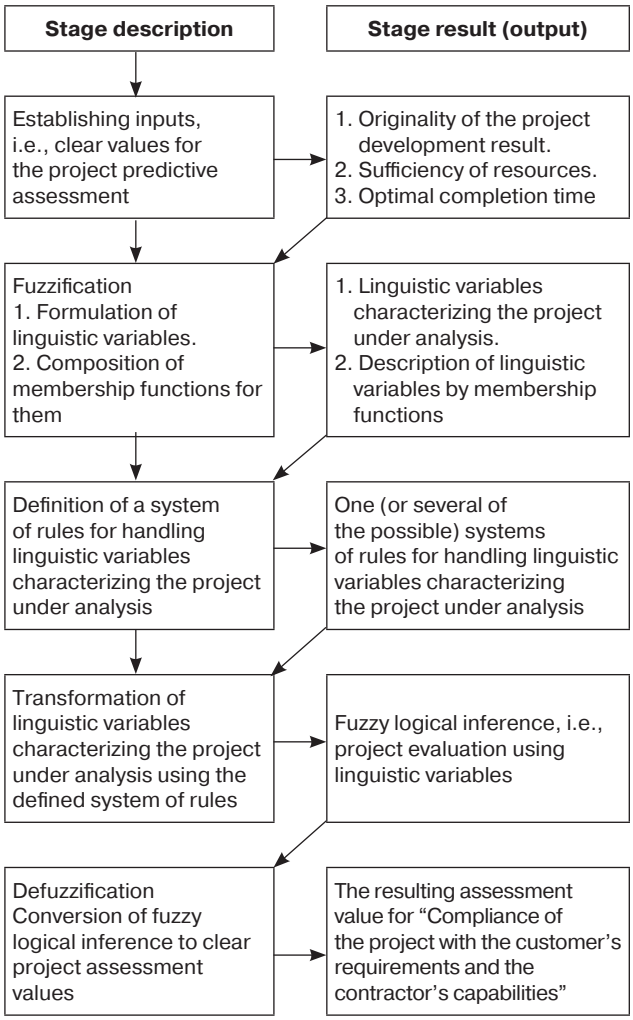


Fig. 1. Stages of the method for assessing the compliance of the proposed project with the customer's requirements and the contractor's capabilities

1. Problem statement: assessment of the project's feasibility and obtaining original scientific, technical or other results that satisfy the customer. Definition and establishment of inputs: clear values that characterize the contractor's activities and are necessary for their assessment using fuzzy logic.

[1] Dadone P. *Design Optimization of Fuzzy Logic Systems*: Ph. Dissertation. Blacksburg, Virginia; 2001. 197 p.

2. Fuzzification: the conversion of clear input values to fuzzy linguistic variables representing the predicted degree of originality of the project result, the level of resource sufficiency and the optimality of work planning in terms of project stage deadlines. Formulation of linguistic variables and composition of membership function(s) for them. Description of linguistic variables by membership functions.
3. Selection of a system of rules for handling linguistic variables in order to formulate a fuzzy logical conclusion.
4. Transformation of linguistic variables using a system of rules to obtain a fuzzy logical conclusion.
5. Defuzzification, which is the reverse transition from fuzzy logical inference to an output parameter, i.e., a generalized assessment of the planned project.
6. Obtaining the output, which is the degree to which the project complies with the customer's requirements and the contractor's capabilities, thereby satisfying both parties.

One of the parameters of a project is the uniqueness of the result obtained, its novelty, and exceptional originality (it has no analogues or surpasses existing solutions by offering fundamentally new ways of meeting the customer's requirements). The method uses the ratio of the number of original technical, scientific or other solutions to their total number in the project as a unit of measurement of project uniqueness. This value is represented by the letter α and lies within a set of possible values A: $\alpha \in A$. The corresponding linguistic variable is "Project uniqueness."

The second project parameter, β, represents the planned time interval for project implementation and its distribution across project stages. The input data is the ratio of the deadlines requested by the potential project contractor and those set by the project customer. The gradations for assessing compliance with the linguistic variable requirements are different degrees of sufficiency or insufficiency of time for project completion. The linguistic parameter "Sufficiency of time for project completion" is assessed using a scale with values from 0.0 to 1.0, where 0.0 represents a complete lack of compliance with the established deadlines and 1.0 represents full compliance with the contractors' and customer's needs.

The third clear input value, γ, is the ratio of the resources (in terms of volume and type) requested by the prospective project contractor and allocated by the project customer. The linguistic parameter "Resource sufficiency" establishes a scale for the fuzzy assessment of project resource provision and the optimal distribution of resources across project stages. The unit of measurement used to assess compliance with requirements is the level of resource sufficiency for project implementation, which is assessed on a scale from 0.0 to 1.0. Here, 0.0 represents complete insufficiency, while 1.0 represents full compliance with requirements.

The fuzzy logic method is employed to reconcile the assessments of the first and second parties during their interaction, thereby facilitating the attainment of a more objective and mutually acceptable assessment. The method requires three input parameters, which are clear values assessed by the customer and the prospective project contractor. This task is solved by fuzzification, i.e., converting the clear input values into linguistic variables. These variables combine the parameters into a set of fuzzy sets whose boundaries do not have clear, unambiguous values that are perceived and accepted by the interacting parties.

Thus, the uniqueness, originality, and novelty of a project may correspond to varying degrees with the requirements and assessments of the customer and contractor. This is because each party has different requirements for assessing the same parameters, project characteristics, and measures to ensure uniform measurements of the project's main characteristics. In an effort to win the contract, the prospective contractor often overestimates their capabilities and the level of the proposed project development, relying on previous experience of performing work that does not fully meet the requirements of the new project. Conversely, the customer, in an effort to minimize the risk of failing to achieve the project goals, is generally more demanding in terms of the planned project results and its key performance indicators. Therefore, in the fuzzy logic approach, there are two subsets: "high compliance" and "low compliance" with the requirements for the main project parameters. These reflect the positions of both the customer and the potential contractor.

There are no clear and unambiguous parameters or limits for indicators such as resource adequacy and the time interval for project implementation. The approach of assessing in terms of "fast/slow," "much/little," "expensive/cheap," and etc. does not bring the positions of the customer and the contractor any closer together when it comes to making decisions on the project.

The problem of combining estimates in the method under consideration can be solved by assigning the linguistic variables to the clear parameter "Uniqueness, originality, and novelty of project results," where α is "Proportion of original solutions in the project out of their total number," β is "Sufficiency of time for project implementation," and γ is "Level of sufficiency of resources for project implementation."

The result of combining the estimates should be the linguistic variable δ, which is defined as "compliance of the project with the customer's requirements and the contractor's capabilities."

The range of possible values in the assessment of project uniqueness corresponds to the fuzzy set A, which is a set of pairs of the form $<\alpha, \varphi_\alpha>$. The grade of membership of each element α of the fuzzy set in set A ($\alpha \in A$) is described by the membership function $\varphi_\alpha(A)$. This function maps each value α to a number within its possible value range. For the "Project uniqueness" parameter, this range is [0.0; 1.0], i.e., $\varphi_\alpha(A)$: $A \rightarrow [0.0; 1.0]$, where each α belongs to set A: $\forall\alpha \rightarrow \in A$. Here, ∀ is the quantifier of universality of properties for all α belonging to set A.

The linguistic variable "Project uniqueness" corresponds to the two extreme values in the range: "high," with a value of "1.0," and "low," with a value of "0.0," from the entire range of possible values of the fuzzy set A. Between these two extremes, there are many other values of the fuzzy set for this project parameter. The membership function $\varphi_\alpha(A)$ shows the extent to which parameter α, i.e., the project uniqueness assessment, has certain properties of subset A, with functions such that when $\alpha = 1.0$, the degree of uniqueness, novelty and originality of the project is maximal, while $\alpha = 0.0$ means that it is minimal. The assessment using fuzzy sets involves assigning a fuzzy value to one of the two subsets, α_{high} and α_{low}, and their respective membership functions, $\varphi_{\alpha high}(A)$ and $\varphi_{\alpha low}(A)$, These functions complement each other, together making up the full range of values of parameter α belonging to the fuzzy set A, as follows:

$$\varphi_{\alpha high}(A) + \varphi_{\alpha low}(A) = 1; \forall\alpha \rightarrow \in A. \quad (1)$$

The innovative nature of project activities consists, among other things, in the fact that the uniqueness, novelty and originality of the development are usually formed at the early stages of the project. Therefore, the membership functions of the "Project uniqueness" parameter in subsets α_h and α_l are best represented by the power function:

$$\varphi_{\alpha high}(A) = \alpha^n \text{ and } \varphi_{\alpha low}(A) = 1 - \alpha^n, \quad (2)$$

where n [0.0; 1.0] is the project innovation index, taking values from 0.0 to 1.0. The value of n indicates the degree of novelty, originality, and patentability of the scientific and technical solutions usually set out in project developments at the early stages, as well as in marketing, the search for and selection of new promising areas of research, and the development of innovative approach.

The membership function graphs for subsets α_h and α_l at the selected range of their possible values, $A \rightarrow [0.0; 1.0]$, for exponents $n = 0.7$, $n = 0.5$, and $n = 0.3$, are shown in Fig. 2.

The area on the graph bounded by the coordinate axes and the membership function line $\varphi_{\alpha high}(A)$ represents the subset of uniqueness with high levels of project novelty ratings, while the membership function line $\varphi_{\alpha low}(A)$ represents the subset with low levels of development originality ratings for which the project is rated as "not unique." The intersection of these subsets represents the subset with a mean estimate for this parameter.

The second project parameter, β, is the planned time interval for project implementation and its division into stages. The membership function of the linguistic variable "Sufficiency of time for project implementation" for the clear parameter "Time interval for project implementation" has the following gradations: "sufficient" and "insufficient." These correspond to the two fuzzy subsets of the total number of fuzzy sets, $\beta \in B$, corresponding to the planned deadlines for implementing project B:

$$\Psi_{\beta high}(B) = \beta^m, \ \Psi_{\beta low}(B) = 1 - \beta^m. \quad (3)$$

The fuzzy subset $\beta_{low}(B)$ corresponds to insufficient time being planned for project completion. The subset $\beta_{high}(B)$ corresponds to a time interval that is sufficient for project completion, as well as an optimal distribution of this time across the project's stages. The maximum value of the membership function "Sufficiency of time for project implementation" is 1.0. The ratio of the membership functions $\Psi_{\beta low}(B)$ and $\Psi_{\beta high}(B)$ for the subsets $\beta_{low}(B)$ and $\beta_{high}(B)$ respectively, is given by:

$$\Psi_{\beta low}(B) + \Psi_{\beta high}(B) = 1; \forall\beta \rightarrow \in B. \quad (4)$$

Given the unevenness and considerable uncertainty in the accuracy of the time distribution across project stages, confidence in the assessment of the parameter "Sufficiency of time for project implementation" belonging to the "sufficient" subset increases nonlinearly as the time interval for project implementation approaches its highest values. Conversely, the grade of membership in the "insufficient" subset also decreases nonlinearly. This type of dependence (Fig. 3) for the subsets $\Psi_{\beta low}(B)$ and $\Psi_{\beta high}(B)$ corresponds to power functions with exponents $m > 1$.

The degree indicator m, being the "fuzziness indicator," represents the level of requirements for assigning a linguistic parameter to the "sufficient" or "insufficient" subsets. The higher its value, the more

the linguistic parameter belongs to a particular subset at higher values of the project completion time β, and, accordingly, the higher the level of certainty and rigor in the assessment of the linguistic parameter. When $m = 0$, i.e., for the linear membership function, the fuzzy set of values of the parameter "Sufficiency of time for project implementation" is divided into two equal fuzzy subsets. As m increases, the fuzzy and crisp estimates tend to converge and become equal at $\beta = 1$.

In the graph (Fig. 3), the area bounded by the coordinate axes and the lines $\Psi_{\beta high}(B)$ and $\Psi_{\beta low}(B)$ is a subset for which the time interval for project completion is estimated to be sufficient or insufficient, respectively.

The intersection of these two areas represents the mean estimate of the sufficiency of time for project implementation.

The third clear input value, γ, is the ratio of the volume and types of resources requested by the prospective project contractor and allocated by the project customer. The membership function of the linguistic variable "sufficiency of resources" according to the "Ratio of resources" parameter (volume and types) requested by the prospective project contractor and allocated by the project customer has two gradations: "sufficient" and "insufficient."

These correspond to two fuzzy subsets of the total number of fuzzy sets $\gamma \in \Gamma$, which corresponds to the ratio of resources Γ requested by the prospective project contractor and allocated by the customer. Assessments of resource provision as "insufficient" for the performance of project work are represented by the fuzzy subset $\chi_{\gamma low}(\Gamma)$; on the other hand, assessments of the sufficiency of resources for the performance of project work are represented by the subset $\chi_{\gamma high}(\Gamma)$, as follows:

$$\chi_{\gamma low}(\Gamma) + \chi_{\gamma high}(\Gamma) = 1; \forall\gamma \rightarrow \in \Gamma. \quad (5)$$

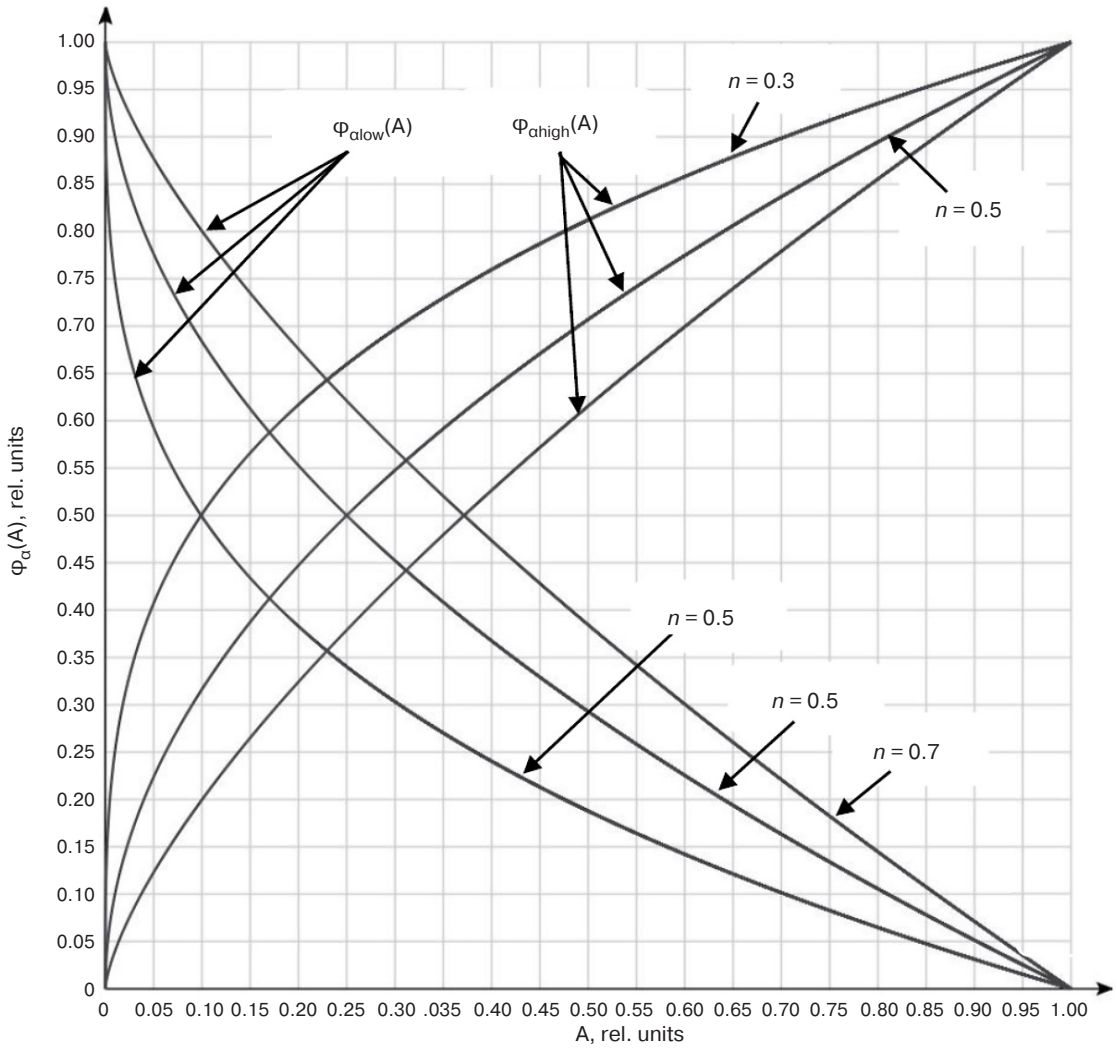


Fig. 2. Membership functions $\varphi_\alpha(A)$ of the linguistic variable "Project uniqueness" in subsets $\varphi_{\alpha high}(A)$ being "unique" and $\varphi_{\alpha low}(A)$ being "not unique" for various indicators of project innovativeness n

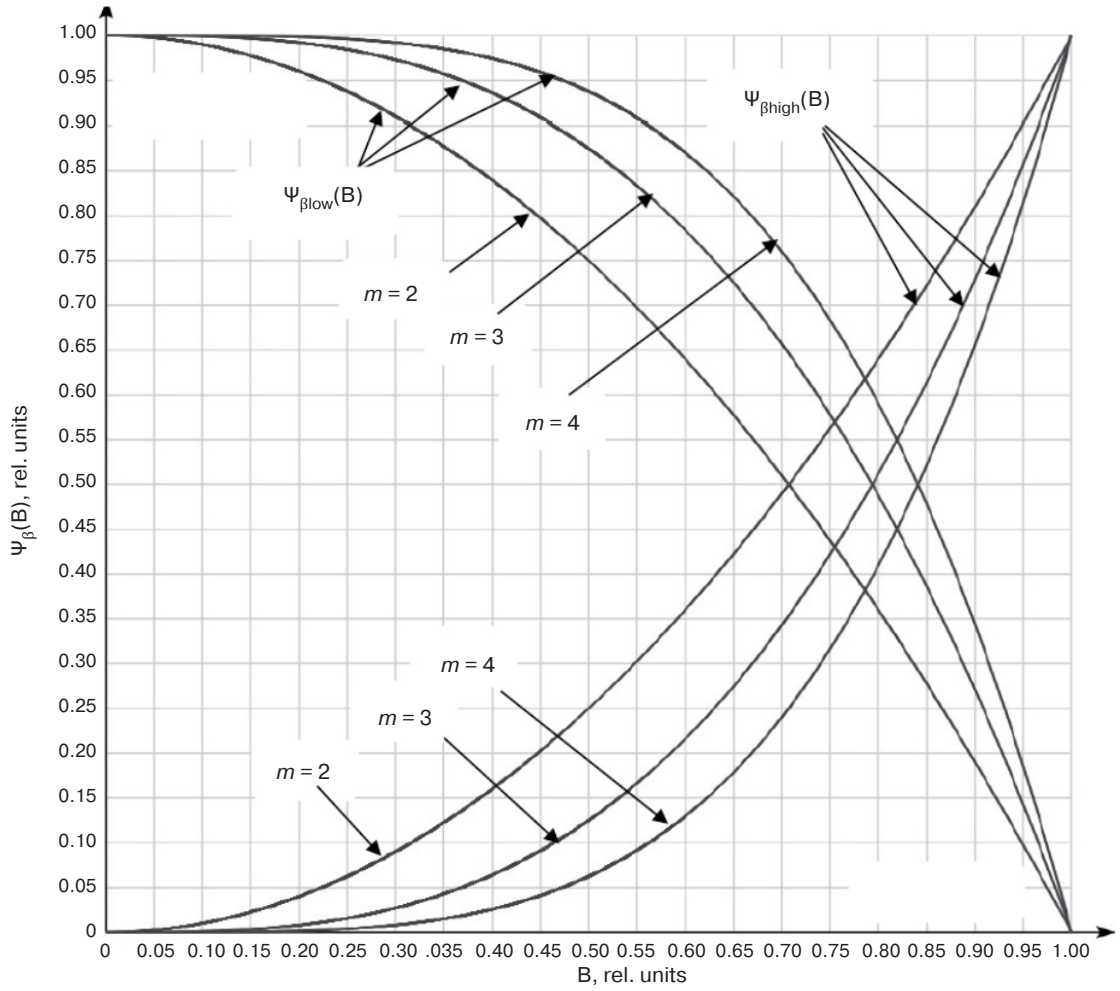


Fig. 3. Membership functions $\Psi_I(B)$ of the linguistic variable "Sufficiency of time for project implementation" in subsets $\Psi_{\beta high}(B)$ being "sufficient" and $\Psi_{\beta low}(B)$ being "insufficient" for various indicators of the degree of certainty of estimates *m*

A distinctive feature of providing material resources for a project is the uneven distribution of these resources across the stages of the project. The creative potential and intellectual resources of contractors are used at the initial stages of project activity, while the main consumption of material resources is concentrated at the stages of development and research, modelling, prototyping, measurements, and testing of prototypes and other results of project activity. At the final stages, which involve documentation and acceptance, the project is supported by information along with organizational and management resources. The distribution of material and intellectual resources across the project stages is best described by the membership function of the linguistic variable "Sufficiency of resources" in sigmoid form (Fig. 4), as follows:

$$\chi_{\gamma high}(\Gamma) = 1 - e^{-2k\gamma}; \chi_{\gamma low}(\Gamma) = e^{-2k\gamma}, \quad (6)$$

where *k* is the indicator of the increase in the need for material resources at different stages of the project.

The linguistic parameter δ, "Compliance of the project with the customer's requirements and the contractor's capabilities," is evaluated on a three-point scale: "does not comply," "partially complies," and "complies." This parameter is a combination of the assessments of the three linguistic variables discussed above: "Originality of the project," "Sufficiency of resources," and "Sufficiency of time for project implementation."

The parameter "Compliance of the project with the customer's requirements and the contractor's capabilities" incorporates two assessments from each of the parties involved: the project customer organization and the contractor organization. The incorporation of these assessments increases the objectivity of the final assessment.

In order to analyze the linguistic variable "Compliance of the project with the customer's requirements and the contractor's capabilities" in the presented fuzzy evaluation model, the following system of rules, incorporating the logical operators "AND" and "OR," is selected from the possible options [15].

The logical operator "AND" denotes the intersection of three fuzzy sets with the intersection membership function:

$$\delta_{A\cap B\cap\Gamma} = \min(\varphi_\alpha(A);\psi_\beta(B);\chi_\gamma(\Gamma)). \quad (7)$$

The logical operator "OR" represents the union of three fuzzy sets. The membership function in this union is as follows:

$$\delta_{A\cap B\cap\Gamma} = \max(\varphi_\alpha(A);\psi_\beta(B);\chi_\gamma(\Gamma)). \quad (8)$$

A system of three rules is applied to defuzzify the fuzzy assessment based on these three linguistic variables and obtain an evaluation of the generalized indicator "Compliance of the project with the customer's requirements and the contractor's capabilities." The first rule is:

R_1: if $(\alpha \in \alpha_{high})$ AND $(\beta \in \beta_{high})$ AND $(\gamma \in \gamma_{high})$, then $\Delta \in \delta_{high}$.

This means that if the assessments of all three linguistic parameters ("Project uniqueness," "Sufficiency of time," and "Sufficiency of material resources") belong to subsets with a "high" compliance, then the assessment of the "Compliance of the project with the customer's requirements and the contractor's capabilities" indicator, Δ, belongs to the "high" subset.

The second rule is:

R_2: if $(\alpha \in \alpha_{low})$ AND $(\beta \in \beta_{low})$ AND $(\gamma \in \gamma_{low})$, then $\Delta \in \delta_{low}$.

This means that if all three assessments of the project parameters in terms of uniqueness, sufficiency of time and resources for project implementation, and compliance with requirements belong to the "low" subsets, then the compliance assessment of the "Compliance of the project with the customer's requirements and the contractor's capabilities" indicator Δ belongs to the "low" subset.

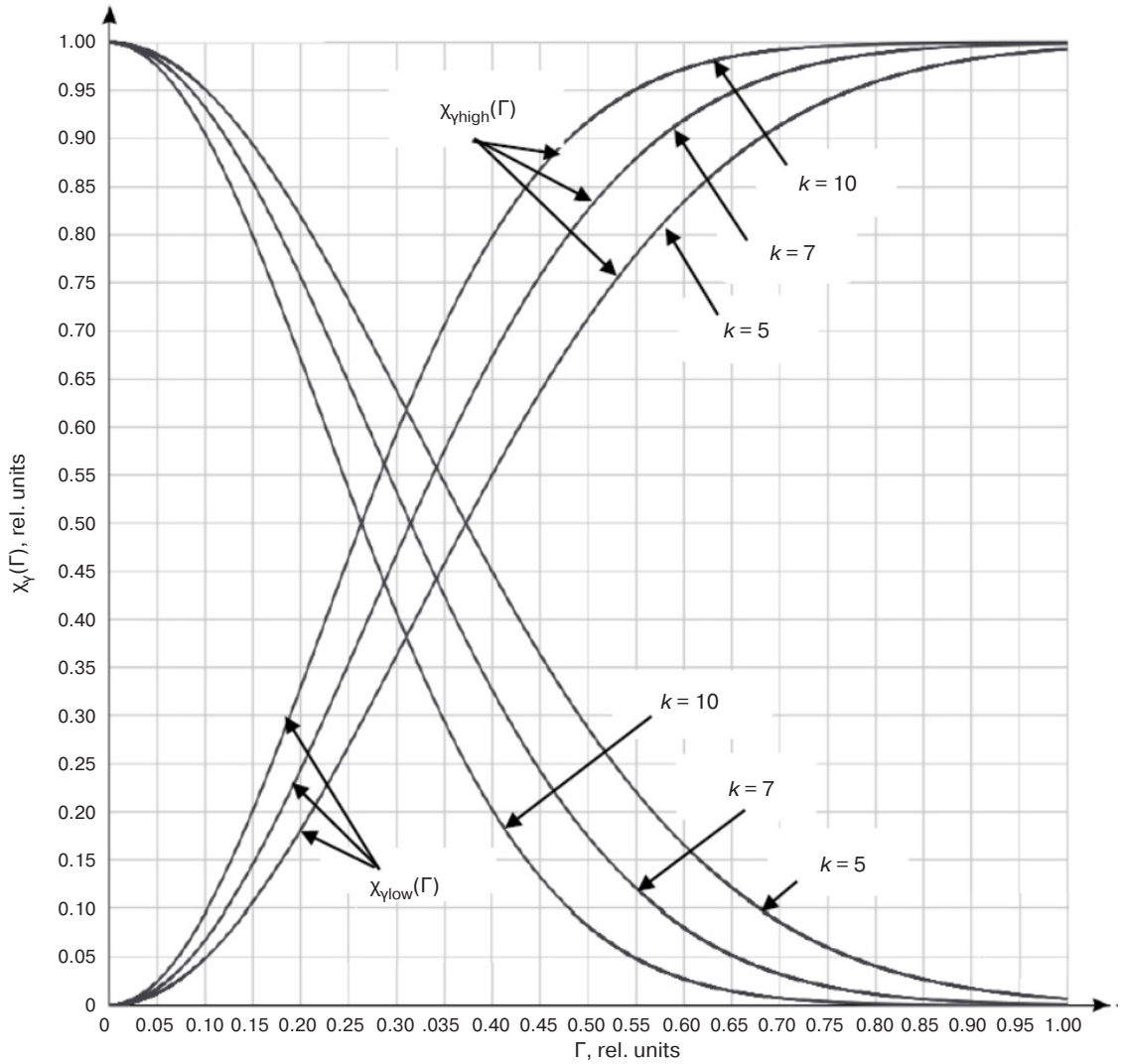


Fig. 4. Membership functions of the linguistic variable "Sufficiency of resources" for different values of the indicator *k* of increased demand for material resources at project stages

It follows from rule R_3 that:

if ($\alpha \in \alpha_{high}$) AND ($\beta \in \beta_{high}$) AND ($\gamma \in \gamma_{low}$) OR
OR ($\alpha \in \alpha_{low}$) AND ($\beta \in \beta_{high}$) AND ($\gamma \in \gamma_{high}$) OR
OR ($\alpha \in \alpha_{high}$) AND ($\beta \in \beta_{low}$) AND ($\gamma \in \gamma_{high}$),
then: $\Delta \in \delta_m$.

This means that, if the assessment of at least one of the three linguistic parameters of the project belongs to the "low" subset and the other two belong to the "high" subset, the "Compliance of the project with the customer's requirements and the contractor's capabilities" parameter, Δ, belongs to the "mean" subset.

The fourth rule considers another possible combination of the project's fuzzy linguistic parameter values:

R_4: if ($\alpha \in \alpha_{high}$) AND ($\beta \in \beta_{low}$) AND ($\gamma \in \gamma_{low}$) OR
OR ($\alpha \in \alpha_{low}$) AND ($\beta \in \beta_{low}$) AND ($\gamma \in \gamma_{high}$) OR
OR ($\alpha \in \alpha_{low}$) AND ($\beta \in \beta_{high}$) AND ($\gamma \in \gamma_{low}$),
then: $\Delta \in \delta_{low}$.

This means that if one of the three linguistic parameters of the project belongs to the subset with a high compliance, and the other two belong to subsets with a low compliance, then the "Compliance of the project with the customer's requirements and the contractor's capabilities" parameter, Δ, belongs to the subset with a low compliance, or "does not comply."

The conversion of fuzzy logical conclusions according to established rules to clear values corresponding to the management system assessment (defuzzification) provides an assessment of the parameter "Compliance of the project with customer requirements and contractor capabilities," taking into account the assessments of both parties involved in the project.

For example, let us consider the assessment of this parameter based on input data, i.e., the expert assessment of three parameters of the planned project based on the contractor's and customer's previous experience.

Thus, if the input data is an expert (i.e., clear) assessment of the uniqueness of the project with a value of 0.75, then the membership function value for project uniqueness in the "high" compliance subset, as calculated by formula (2) for a borrowing level of $n = 0.7$, is 0.817. For $n = 0.5$, the value is 0.866, while for $n = 0.3$, the corresponding value is 0.903. Similarly, for the same expert assessment of the possibility of achieving project uniqueness with an index of 0.75, the value of the membership function in the "low" subset is 0.183 for $n = 0.7$, 0.134 for $n = 0.5$, and 0.083 for $n = 0.3$.

For an expert assessment of the input parameter (i.e., the ratio of the time interval requested by the contractor and allocated by the customer for project implementation), $\beta = 0.80$, the value of the "Sufficiency of time for project implementation" membership function in the "sufficient" subset calculated using formula (3) is 0.64 for the "fuzziness index," $m = 2$, 0.51 for $m = 3$, and 0.4 for $m = 4$. For the same input parameter value, $\gamma = 0.80$, the membership function value in the "insufficient" subset is 0.36 for $m = 2$, 0.49 for $m = 3$, and 0.59 for $m = 4$.

The values of the membership function of the linguistic variable "Sufficiency of resources" in the subset "sufficient" for the parameter $\gamma = 0.65$, calculated using formula (4) with different indicators of increased need for material resources at project stage k, are as follows: 0.88 for $k = 5$; 0.93 for $k = 7$; and 0.98 for $k = 10$.

Accordingly, the membership grade of the linguistic variable "Sufficiency of material resources" in the "low" subset is estimated as 0.12 for $k = 5$; 0.07 for $k = 7$; and 0.02 for $k = 10$.

The resulting membership function values for each project parameter show that they all belong to subsets indicating high compliance with the project uniqueness requirements, as well as sufficient time and material resources. The defuzzification and evaluation of the membership of the parameter "Compliance of the project with the customer's requirements and the contractor's capabilities" in the "high" subset is determined by rule R_1 with the logical operator "AND," as follows:

$$\text{for } n = 0.7,\ m = 2,\ k = 5:\ \delta_{A\cap B\cap \Gamma} = \min(\varphi_{high}(A);\psi_{high}(B);\gamma_{high}(\Gamma)) = \min(0.82;0.64;0.88) = 0.64;$$

$$\text{for } n = 0.5,\ m = 3,\ k = 7:\ \delta_{A\cap B\cap \Gamma} = \min(\varphi_{high}(A);\psi_{high}(B);\gamma_{high}(\Gamma)) = \min(0.87;0.51;0.93) = 0.51;$$

$$\text{for } n = 0.3,\ m = 4,\ k = 10:\ \delta_{A\cap B\cap \Gamma} = \min(\varphi_{high}(A);\psi_{high}(B);\gamma_{high}(\Gamma)) = \min(0.90;0.41;0.98) = 0.41.$$

According to rule R_2, the assessment of the membership grade of the linguistic variable "Compliance of the project with the customer's requirements and the contractor's capabilities" in the fuzzy subset "low" yields the following results:

$$\text{for } n = 0.7,\ m = 2,\ k = 5:\ \delta_{A\cap B\cap \Gamma} = \min(\varphi_{low}(A);\psi_{low}(B);\gamma_{low}(\Gamma)) = \min(0.18;0.46;0.12) = 0.12;$$

for $n = 0.5$, $m = 3$, $k = 7$: $\delta_{A\cap B\cap\Gamma} = \min(\varphi_{low}(A);\psi_{low}(B);\gamma_{low}(\Gamma)) = \min(0.13;0.49;0.07) = 0.07;$

for $n = 0.3$, $m = 4$, $k = 10$: $\delta_{A\cap B\cap\Gamma} = \min(\varphi_{low}(A);\psi_{low}(B);\gamma_{low}(\Gamma)) = \min(0.10;0.59;0.02) = 0.02.$

This means that the membership grade of the parameter "Compliance of the project with the customer's requirements and the contractor's capabilities" is higher in the "high" subset than in the "low" subset. For the "high" subset, the estimated grades of membership for various factors affecting project implementation are 0.64, 0.51, and 0.41 points, while for the "low" subset they are 0.12, 0.07, and 0.02 points under the same conditions.

The membership of the indicator "Compliance of the project with the customer's requirements and the contractor's capabilities" in the subset "mean" is determined by rule R_3 using the following ratios:

for $n = 0.7$, $m = 2$, $k = 5$:

$$\delta_{A\cap B\cap\Gamma} = \max(\min(\varphi_{low}(0.75);\psi_{high}(0.80);\gamma_{high}(0.65));\min(\varphi_{high}(0.75);\psi_{low}(0.80);\gamma_{high}(0.65));$$
$$\min(\varphi_{high}(0.75);\psi_{high}(0.80);\gamma_{low}(0.65))) = \max(\min(0.18;0.64;0.88);$$
$$\min(0.87;0.64;0.93);\min(0.82;0.46;0.12)) = \max(0.18;0.46;0.12) = \mathbf{0.46};$$

for $n = 0.5$, $m = 3$, $k = 7$:

$$\delta_{A\cap B\cap\Gamma} = \max(\min(\varphi_{low}(0.75);\psi_{high}(0.80);\gamma_{high}(0.65));\min(\varphi_{high}(0.75);\psi_{low}(0.80);\gamma_{high}(0.65));$$
$$\min(\varphi_{high}(0.75);\psi_{high}(0.80);\gamma_{low}(0.65))) = \max(\min(0.13;0.51;0.93);$$
$$\min(0.82;0.49;0.88);\min(0.87;0.51;0.07)) = \max(0.13;0.49;0.07) = \mathbf{0.49};$$

for $n = 0.3$, $m = 4$, $k = 10$:

$$\delta_{A\cap B\cap\Gamma} = \max(\min(\varphi_{low}(0.75);\psi_{high}(0.80);\gamma_{high}(0.65));\min(\varphi_{high}(0.75);\psi_{low}(0.80);\gamma_{high}(0.65));$$
$$\min(\varphi_{high}(0.75);\psi_{high}(0.80);\gamma_{low}(0.65))) = \max(\min(0.10;0.41;0.98);$$
$$\min(0.90;0.59;0.98);\min(0.90;0.41;0.02)) = \max(0.10;0.59;0.02) = \mathbf{0.59}.$$

The estimated values of the membership functions of the parameter "Compliance of the project with the customer's requirements and the contractor's capabilities" in the "high," "mean," and "low" subsets, when calculated according to rules R_1, R_2, and R_3, yield the following results:

for $n = 0.7$, $m = 4$, $k = 10$: 0.64; **0.46**; 0.12;
for $n = 0.5$, $m = 3$, $k = 7$: 0.51; **0.49**; 0.07;
for $n = 0.3$, $m = 4$, $k = 10$: 0.41; **0.59**; 0.02.

By comparing the values obtained for the degree of agreement between the customer and contractor's assessments of the planned project the following output is achieved:

- the customer decides on the prospects and feasibility of setting up work on the project, taking various influencing factors into account, and to review and change the project requirements, the allocated resources, and the deadlines;
- the contractor identifies additional opportunities to implement the project under the conditions proposed by the customer.

2. METHODOLOGY FOR ASSESSING THE COMPLIANCE OF THE PROJECT WITH THE CUSTOMER'S REQUIREMENTS AND CONTRACTOR'S CAPABILITIES USING AN APPROACH BASED ON FUZZY SET THEORY

The application of the above method for assessing the compliance of the project with the customer's requirements and the contractor's capabilities, while taking into account the perceptions of both interacting parties, consists of a sequence of stages which are presented in Fig. 1[2] [15, 16] that comprises:

- identifying the project, its goals and purpose, and the planned results, as well as describing the project's most significant and essential properties, parameters and characteristics (in particular, its uniqueness, scope and types of resources, and time frame);

[2] Dadone P. *Design Optimization of Fuzzy Logic Systems*: Ph. Dissertation. Blacksburg, Virginia; 2001. 197 p.

- formulating the task of obtaining a joint assessment of the project's compliance with the customer's requirements and the contractor's capabilities, which will serve as the basis for decisions regarding the project's feasibility and advisability;
- determining the values of the input parameters of the modelled project using calculations and/or expert assessment, and establishing them for further fuzzification;
- fuzzification, i.e., the transition to fuzziness and the conversion of clear input values (the predicted degree of originality of the project result, the level of resource sufficiency, and the project completion dates) into fuzzy linguistic variables;
- formulating linguistic variables and defining their membership functions; describing linguistic variables using membership functions;
- selecting a system of rules for handling linguistic variables in order to formulate fuzzy logical conclusions.
- transforming linguistic variables using a system of rules to obtain fuzzy logical conclusions;
- defuzzification – reverse transition from a fuzzy logical conclusion to a clear value (the output parameter), i.e., a generalized assessment of whether the planned project complies with the customer's requirements and the contractor's capabilities;
- obtaining an output, i.e., a quantitative value of the project's compliance with the customer's requirements and the contractor's capabilities, with acceptance by both parties.

When supplemented with requirements for contractor competence, the documented methodology can serve as a tool for making and implementing decisions regarding the management of project activities, including research and development work [12–14].

CONCLUSIONS

The presented method and technique allow projects to be evaluated, whether they are planned for implementation or have already been completed, based on the results obtained. The evaluation of project management's individual stages using the presented method will also improve its effectiveness, as will further detailing, breaking down, and evaluating individual types of work if necessary.

Along with the corresponding methodology for its application in adopting and implementing decisions based on the generalized assessment obtained, the application of this analytical method, which represents an alternative to expert methods, will ensure greater validity, objectivity and effectiveness.

The documented evaluation procedure ensures reproducible results based on the identity and repeatability of actions by different contractors. This reduces dependence on subjective qualities and increases the objectivity of the evaluation, as well as ensuring comparability with previous evaluation results for similar projects. The use of consistent criteria and procedures also permits a comparison of the results of project activity assessments by different contractors.

A promising development for the method and its associated methodology would be to apply membership functions to the analysis of fuzzy systems. This would take into account a greater number of characteristics of projects and their components, as well as determining the potential of contractors and customer requirements. The increasing volume of data and the resulting complexity of computational processes will require new resources to process information for both assessment and solution development. One effective solution to this problem could be to apply digital and artificial intelligence technologies to the processes of assessment and development of management decisions.

Due to the invariance of the project activity attributes criteria, the presented method and techniques recommend themselves as a universal tool for managing various types of projects and activities, including research and development work.

REFERENCES

1. Tsipes G.L. Methods of an assessment of efficiency of project-oriented activity. Review of current state and prospects of development. *Upravlenie proektami i programmami.* 2009;03(19):190–205 (in Russ.).
2. Galushko A.D., Yurina V.S., Yurlova I.S. Review of methods of assessment and analysis of investment projects. *NovaInfo.* 2016;3(57):317–328 (in Russ.).
3. Gracheva M.V. (Ed.). *Risk-analiz investitsionnogo proekta* (*Risk Analysis of Investment Project*). Moscow: YuNITI; 2001. 350 p. (in Russ.). ISBN 5-238-00292-0
4. Illarionov A.V., Klimenko E.Yu. *Portfel' proektov: Instrument strategicheskogo upravleniya predpriyatiem* (*Project Portfolio: Tool for Strategic Enterprise Management*). Moscow: Alpina Publisher; 2013. 312 p. (in Russ.). ISBN 978-5-9614-2261-0
5. Razu M.L. *Upravlenie proektom: osnovy proektnogo upravleniya* (*Project Management: Fundamentals of Project Management*). Moscow: KnoRus; 2015. 756 p. (in Russ.). ISBN 978-5-406-04370-7

6. Gracheva M.V. *Proektnyi analiz: uchet riskov* (*Project Analysis: Risk Accounting*). Moscow: Prospekt; 2017. 178 p. (in Russ.). ISBN 978-5-392-26095-9
7. Turner J.R. *The Handbook of Project-Based Management*. London, NY: McGrow-Hill Book; 1995. 540 p.
[Turner J.R. *Rukovodstvo po proektno-orientirovannomu upravleniyu* (*Handbook of Project-Oriented Management*): transl. from Engl. Moscow: Grebennikov Publishing House; 2007. 552 p. (in Russ.). ISBN 5-93890-027-1]
8. Benko C., Mak-Farlan F.W. *Connecting the Dots: Aligning Projects with Objectives in Unpredictable Times*. Harvard Business Review Press; 2003. 246 p.]
[Benko C., Mak-Farlan F.W. *Upravlenie portfelyami proektov: sootvetstvie proektov strategicheskim tselyam kompanii* (*Connecting the Dots: Aligning Projects with Objectives in Unpredictable Times*): transl. from Engl. Moscow: Vil'yams; 2007. 240 p. (in Russ.). ISBN 978-5-8459-1059-2]
9. Mazur I.I., Shapiro V.D., Olderogge N.G. *Upravlenie proektami* (*Project Management*). Moscow: Omega-L; 2001. 664 p. (in Russ.). ISBN 5-370-00049-2
10. Kovalev V.V. *Metody otsenki investitsionnykh proektov* (*Methods of Investment Project Evaluation*). Moscow: Finansy i statistika; 2008. 144 p. (in Russ.).
11. Tyapukhin A.P. On the issue of the concept of value chain management. *Upravlencheskoe konsul'tirovanie = Administrative Consulting*. 2023;11:46–59 (in Russ.). https://doi.org/10.22394/1726-1139-2023-11-46-59
12. Sidorin V.V., Khalilyulina N.B. Model and method of evaluating the effectiveness of the project. *Fundamental'nye problemy tekhniki i tekhnologii = Fundamental and Applied Problems of Engineering and Technology*. 2024;5(367):81–90 (in Russ.).
13. Sidorin V.V. Design and development of radio-electronic devices using artificial intelligence technologies. In: *Actual Problems and Prospects of Development of Radio Engineering and Infocommunication Systems* (*RADIOINFOCOM-2024*): *Proceedings of the 8th International Scientific and Practical Conference*. Moscow: RTU MIREA; 2024. P. 566–573 (in Russ.).
14. Sidorin V.V., Khalilyulina N.B. Forecasting and management of research and development work by Markov methods. *Vestnik Sankt-Peterburgskogo gosudarstvennogo universiteta tekhnologii i dizaina. Seriya 4. Promyshlennye tekhnologii = Vestnik of St. Petersburg State University of Technology and Design. Series 4. Industrial Technologies*. 2023;3:58–62 (in Russ.). https://www.elibrary.ru/vgtnor
15. Piegat A. *Fuzzy Modeling and Control*. NY: Physica Heidelberg; 2001. 728 p.
[Piegat A. *Nechetkoe modelirovanie i upravlenie* (*Fuzzy Modeling and Control*): transl. from Engl. Moscow: Binom. Laboratoriya znanii; 2009. 798 p. (in Russ.). ISBN 978-5-94774-353-1]
16. Kandel A. *Fuzzy Control Systems.* Boca Raton: CRC Press; 1994. 656 p.
17. Ledeneva T.M., Reshetnikov A.D. Aspects of the implementation of the fuzzy inference mechanism in fuzzy systems. *Mezhdunarodnyi nauchno-issledovatel'skii zhurnal = International Scientific Research Journal*. 2021;6(108):107–117 (in Russ.). https://doi.org/10.23670/IRJ.2021.108.6.018
18. Kerzner H. *Strategic Planning for Project Management Using a Project Management Maturity Model.* NY: J. Wiley; 2001. 280 p.
[Kerzner H. *Strategicheskoe planirovanie dlya upravleniya proektami s ispol'zovaniem modeli zrelosti* (*Strategic Planning for Project Management Using a Project Management Maturity Model*): transl. from Engl. Moscow: AITi; DMK Press; 2003. 318 p. (in Russ.). ISBN 5-94074-211-4]
19. Guseva E.G. *Monitoring i otsenka proektov* (*Monitoring and Evaluation of Projects*). St. Petersburg: TsRNO; 2014. 43 p. (in Russ.).

СПИСОК ЛИТЕРАТУРЫ

1. Ципес Г.Л. Методы оценки эффективности проектно-ориентированной деятельности. Обзор текущего состояния и перспектив развития. *Управление проектами и программами*. 2009;03(19):190–205.
2. Галушко А.Д., Юрина В.С., Юрлова И.С. Обзор методов оценки и анализа инвестиционных проектов. *NovaInfo*. 2016;3(57):317–328. https://elibrary.ru/xeukzr
3. *Риск-анализ инвестиционного проекта*; под ред. М.В. Грачёвой. М.: ЮНИТИ; 2001. 350 с. ISBN 5-238-00292-0
4. Илларионов А.В., Клименко Э.Ю. *Портфель проектов: Инструмент стратегического управления предприятием.* М.: Альпина Паблишер; 2013. 312 с. ISBN 978-5-9614-2261-0
5. Разу М.Л. *Управление проектом: основы проектного управления.* М.: КноРус; 2015. 756 с. ISBN 978-5-406-04370-7
6. Грачёва М.В. *Проектный анализ: учет рисков.* М.: Проспект; 2017. 178 с. ISBN 978-5-392-26095-9
7. Тернер Дж.Р. *Руководство по проектно-ориентированному управлению:* пер. с англ. М.: Издательский дом Гребенникова; 2007. 552 с. ISBN 5-93890-027-1
8. Бенко К., Мак-Фарлан Ф.У. *Управление портфелями проектов: соответствие проектов стратегическим целям компании*: пер. с англ. М.: Вильямс; 2007. 240 с. ISBN 978-5-8459-1059-2
9. Мазур И.И., Шапиро В.Д., Ольдерогге Н.Г. *Управление проектами*. М.: Омега-Л; 2001. 664 с. ISBN 5-370-00049-2
10. Ковалев В.В. *Методы оценки инвестиционных проектов*. М.: Финансы и статистика; 2008. 144 с.
11. Тяпухин А.П. К вопросу о концепции управления цепями создания ценностей. *Управленческое консультирование*. 2023;11:46–59. https://doi.org/10.22394/1726-1139-2023-11-46-59

12. Сидорин В.В., Халилюлина Н.Б. Модель и метод оценки результативности проекта. *Фундаментальные проблемы техники и технологии*. 2024;5(367):81–90.
13. Сидорин В.В. Проектирование и разработка радиоэлектронных средств с применением технологий искусственного интеллекта. В сб.: *Актуальные проблемы и перспективы развития радиотехнических и инфокоммуникационных систем* («*Радиоинфоком-2024*»): *Сборник научных статей по материалам VIII Международной научно-практической конференции*. М.: РТУ МИРЭА; 2024. С. 566–573.
14. Сидорин В.В., Халилюлина Н.Б. Прогнозирование и управление выполнением научно-исследовательских и опытно-конструкторских работ Марковскими методами. *Вестник Санкт-Петербургского государственного университета технологии и дизайна. Серия 4. Промышленные технологии*. 2023;3:58–62. https://www.elibrary.ru/vgtnor
15. Пегат А. *Нечеткое моделирование и управление*: пер. с англ. М.: Бином. Лаборатория знаний; 2009. 798 с. ISBN 978-5-94774-353-1
16. Kandel A. *Fuzzy Control Systems.* Boca Raton: CRC Press; 1994. 656 p.
17. Леденёва Т.М., Решетников А.Д. Особенности реализации механизма нечеткого логического вывода в нечетких системах. *Международный научно-исследовательский журнал*. 2021;6(108):107–117. https://doi.org/10.23670/IRJ.2021.108.6.018
18. Керцнер Г. *Стратегическое планирование для управления проектами с использованием модели зрелости:* пер. с англ. М.: АйТи; ДМК Пресс; 2003. 318 с. ISBN 5-94074-211-4
19. Гусева Е.Г. *Мониторинг и оценка проектов*. СПб.: ЦРНО; 2014. 43 с.

About the Author

Viktor V. Sidorin, Dr. Sci. (Eng.), Professor, Institute for Testing and Certification of Weapons and Military Equipment (10, Ehlektrodnaya ul., Moscow, 111524 Russia). E-mail: wwsid@yandex.ru. RSCI SPIN-code 1786-8245, https://orcid.org/0009-0003-3555-4687

Об авторе

Сидорин Виктор Викторович, д.т.н., профессор, Автономная некоммерческая организация «Институт испытаний и сертификации вооружения и военной техники» (АНО «ИнИС ВВТ») (111524, Россия, Москва, Электродная ул., д. 10). E-mail: wwsid@yandex.ru. SPIN-код РИНЦ 1786-8245, https://orcid.org/0009-0003-3555-4687

Translated from Russian into English by K. Nazarov
Edited for English language and spelling by Thomas A. Beavitt