THE MINISTRY OF SCIENCE AND HIGHER EDUCATION OF THE RUSSIAN FEDERATION

ISSN 2687-0517

Computing, Telecommunications and Control

Vol. 18, No. 4
2025

Peter the Great St. Petersburg
Polytechnic University
2025

The journal is included in the List of Leading PeerReviewed Scientific Journals and other editions to publish major findings of PhD theses for the research degrees of Doctor of Sciences and Candidate of Sciences.

Open access journal is to publish articles of a high scientific level covering advanced experience, research results, theoretical and practical problems of informatics, electronics, telecommunications, and control.

The journal is indexed by Ulrich's Periodicals Directory, Google Scholar, EBSCO, ProQuest, Index Copernicus, VINITI RAS Abstract Journal (Referativnyi Zhurnal), VINITI RAS Scientific and Technical Literature Collection, Russian Science Citation Index (RSCI) database Scientific Electronic Library and Math-Net.ru databases.

The journal is registered with the Federal Service for Supervision in the Sphere of Telecom, Information Technologies and Mass Communications (ROSKOMNADZOR). Certificate ЭЛ No. ФС77-77378 issued 25.12.2019.

Editorial office

Dr. Sc., Professor A.S. Korotkov – Editor-in-Chief;

Ph.Ch.S. Bastian – literary editor, proofreader; G.A. Pyshkina – editorial manager; A.A. Kononova – computer layout; I.E. Lebedeva – English translation.

Address: 195251 Polytekhnicheskaya Str. 29, St. Petersburg, Russia.

+7 (812) 552-6216, e-mail: infocom@spbstu.ru

Release date: 30.12.2025

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ISSN 2687-0517

Информатика, телекоммуникации и управление

Том 18, № 4
2025

Санкт-Петербургский политехнический
университет Петра Великого
2025

ИНФОРМАТИКА, ТЕЛЕКОММУНИКАЦИИ И УПРАВЛЕНИЕ

Журнал с 2002 года входит в Перечень ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные результаты диссертаций на соискание ученой степени доктора и кандидата наук.

Сетевое издание открытого доступа публикует статьи высокого научного уровня, освещающие передовой опыт, результаты НИР, теоретические и практические проблемы информатики, электроники, телекоммуникаций, управления.

Сведения о публикациях представлены в Реферативном журнале ВИНИТИ РАН, в международной справочной системе «Ulrich`s Periodical Directory», в Российской государственной библиотеке. В базах данных: Российский индекс научного цитирования (РИНЦ), Google Scholar, EBSCO, Math-Net.Ru, ProQuest, Index Copernicus.

Журнал зарегистрирован Федеральной службой по надзору в сфере информационных технологий и массовых коммуникаций (Роскомнадзор). Свидетельство о регистрации ЭЛ № ФС77-77378 от 25.12.2019.

Учредитель и издатель: Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Российская Федерация.

Редакция журнала

д-р техн. наук, профессор А.С. Коротков – главный редактор;

Ф.К.С. Бастиан – литературный редактор, корректор; Г.А. Пышкина – ответственный секретарь, выпускающий редактор;

А.А. Кононова – компьютерная вёрстка; И.Е. Лебедева – перевод на английский язык.

Адрес редакции: Россия, 195251, Санкт-Петербург, ул. Политехническая, д. 29.

Тел. редакции +7(812) 552-62-16, e-mail редакции: infocom@spbstu.ru

Дата выхода: 30.12.2025

Contents

Содержание

Computing, Telecommunication and Control, 2025, Vol. 18, No. 4, Pp. 7–19.
Информатика, телекоммуникации и управление. 2025. Том 18, № 4. С. 7–19.

Intelligent Systems and Technologies, Artificial Intelligence

Интеллектуальные системы и технологии, искусственный интеллект

Research article
DOI: https://doi.org/10.18721/JCSTCS.18401
UDC 004.85

CONCEPT-BASED LEARNING IN HETEROGENEOUS TREATMENT EFFECT

L.V. Utkin ✉ , *A.V. Konstantinov* , *N.M. Verbova*

Peter the Great St. Petersburg Polytechnic University,
St. Petersburg, Russian Federation

✉ lev.utkin@gmail.com

Abstract. Estimating Heterogeneous Treatment Effects (HTE) is crucial for personalized decision-making in medicine, economics and engineering. While machine learning models for Conditional Average Treatment Effect (CATE) estimation have become increasingly accurate, they often remain black boxes, providing little insight into why treatments affect individuals differently. This paper introduces CATE-Concept Bottleneck Model (CATE-CBM), a novel framework that integrates concept-based learning with CATE estimation to bridge this interpretability gap. Our approach enforces a concept bottleneck that forces the model to express treatment effects through understandable concepts, enabling transparent reasoning about which concepts drive heterogeneous effects. Through experiments on a modified MNIST dataset, we demonstrate that CATE-CBM maintains competitive accuracy while providing meaningful concept-based explanations of treatment effect heterogeneity. The model successfully identifies how both the presence and absence of specific concepts influence treatment outcomes, offering clinicians and engineers both accurate effect estimates and interpretable rationales for personalized interventions. This work represents the first unification of concept bottleneck models with causal effect estimation, advancing the frontier of explainable artificial intelligence in causal inference.

Keywords: machine learning, concept-based learning, conditional average treatment effect, interpretation, neural network

Acknowledgements: The research was financially supported by the Russian Science Foundation, project "Machine learning models for assessing treatment effect with heterogeneous diagnostic information using expert rules" (Agreement No. 25-11-00021; available online: https://rscf.ru/project/25-11-00021/).

Citation: Utkin L.V., Konstantinov A.V., Verbova N.M. Concept-Based Learning in Heterogeneous Treatment Effect. Computing, Telecommunications and Control, 2025, Vol. 18, No. 4, Pp. 7–19. DOI: 10.18721/JCSTCS.18401

Научная статья
DOI: https://doi.org/10.18721/JCSTCS.18401
УДК 004.85

ОБУЧЕНИЕ НА ОСНОВЕ КОНЦЕПТОВ ДЛЯ ОЦЕНКИ УСЛОВНОГО ЭФФЕКТА ЛЕЧЕНИЯ

Л.В. Уткин ✉, *А.В. Константинов*, *Н.М. Вербова*

Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Российская Федерация

✉ lev.utkin@gmail.com

Аннотация. Оценка условного эффекта лечения имеет решающее значение для персонализированного принятия решений в медицине, экономике и технике. Хотя модели машинного обучения для оценки условного среднего эффекта лечения (CATE) становятся все более точными, они часто остаются «черными ящиками», не давая понимания того, почему лечение по-разному влияет на различных людей. Данная работа представляет новую модель CATE-CBM, которая интегрирует обучение на основе концептов с оценкой CATE, чтобы преодолеть этот разрыв в интерпретируемости. Предлагаемый подход использует обучение на концептах, заставляя модель выражать эффекты лечения через понятные для человека концепты, что позволяет прозрачно объяснять, какие именно концепты обуславливают эффекты. В экспериментах на модифицированном наборе данных MNIST демонстрируется, что CATE-CBM сохраняет конкурентоспособную точность, одновременно предоставляя содержательные объяснения предсказания эффекта лечения на основе концептов. Модель успешно идентифицирует, как присутствие или отсутствие конкретных концептов влияет на результаты лечения, предлагая клиницистам и политикам как точные оценки эффекта, так и интерпретируемые обоснования для персонализированных вмешательств. Данная работа представляет собой первую унификацию моделей с обучением на концептах и оценкой причинно-следственных связей, продвигая границы объяснимого искусственного интеллекта.

Ключевые слова: машинное обучение, обучение на концептах, условный средний эффект лечения, интерпретируемость, нейронная сеть

Финансирование: Исследование выполнено при финансовой поддержке Российского научного фонда в рамках реализации проекта «Модели машинного обучения для оценки эффекта лечения при разнородной диагностической информации с экспертными правилами» (Соглашение № 25-11-00021; https://rscf.ru/project/25-11-00021/).

Для цитирования: Utkin L.V., Konstantinov A.V., Verbova N.M. Concept-Based Learning in Heterogeneous Treatment Effect // Computing, Telecommunications and Control. 2025. Т. 18, № 4. С. 7–19. DOI: 10.18721/JCSTCS.18401

Introduction

The pursuit of explainability and the integration of human-centric reasoning into Machine Learning (ML) has catalyzed the development of Concept-Based Learning (CBL). Unlike conventional models that operate directly on raw, low-level features, CBL utilizes high-level, human-intelligible concepts as intermediate representations for making predictions [1]. This paradigm aims to bridge the gap between data-driven patterns and expert knowledge, leading to models that are not only more interpretable but also more data-efficient and robust [2, 3]. A prominent instantiation of this approach is the Concept Bottleneck Model (CBM), which enforces a compressed, concept-based representation of the input, forcing the final classifier to rely solely on these concepts for prediction [4]. This architectural constraint ensures that the model's decision-making process is intrinsically tied to a vocabulary of meaningful concepts, significantly enhancing its explainability [3].

The application of CBL can extend beyond standard supervised learning into more complex domains, such as estimating Heterogeneous Treatment Effects (HTE). HTE is the recognition that the effect of a treatment (e.g., a drug, a policy) varies across different individuals [5–7]. While HTE describes this general property, the Conditional Average Treatment Effect (CATE) quantifies it by measuring the average treatment effect for a specific subpopulation with given characteristics. To estimate the treatment effect, patients are typically divided into treatment and control groups, and their average outcomes are compared. This comparison provides an estimate of the causal effect of the treatment. Various approaches to estimating this effect are considered in several surveys [5, 8–12].

The primary aim of combining HTE and CBL is to provide explanations for why a certain treatment is predicted to be more effective for one individual than for another. For instance, in a medical context, a model might predict a stronger positive treatment effect for patients characterized by the concepts “high genetic marker expression” and “early disease stage”, providing clinicians with a clear, conceptual rationale for personalized treatment plans. Incorporating CBL into CATE estimation moves us beyond simply knowing that a treatment effect is heterogeneous; it provides the crucial why, explaining this heterogeneity through the lens of understandable concepts. This can lead to more informed and personalized decisions in healthcare, economics and public policy. Moreover, to the best of our knowledge, no existing method combines CBL with HTE or CATE.

Motivated by the above reasoning, we propose a model called CATE-CBM, which integrates concept learning and CATE estimation into a single framework. The model consists of two main components:

1. The first produces concept probabilities, which serve as a type of embedding.
2. The second part solves the CATE estimation problem using these predicted concept probabilities for patients in the treatment and control groups.

An important characteristic of CBL is the interpretation of predictions in terms of human-intelligible concepts. Therefore, this paper demonstrates how to locally interpret the CATE predictions made by our model.

Numerical experiments conducted on a modified MNIST dataset demonstrate how the integration of CATE estimation and CBL can improve accuracy and provide explanations for the CATE-CBM predictions in terms of concepts.

Related work

Concept-based learning

The growing interest in CBL has led to a proliferation of models aimed at improving the interpretability and explainability of ML predictions [1, 2]. These models leverage understandable concepts to make model reasoning more transparent and to align machine decisions with user intuition. The CBM [4] as a special case of CBL serves as a foundational architecture for many CBL approaches. Its efficient two-stage design in predicting concepts from inputs, then targets from concepts, has inspired numerous extensions. These include models that learn continuous concept embeddings [13], probabilistic variants to handle uncertainty [14] and investigations into concept independence and intervention [15, 16]. Further adaptations have integrated powerful pre-trained models like CLIP [17, 18] and addressed performance disparities between different CBM formulations [19].

Survey papers [20–22] comprehensively discuss aspects of CBL, CBM and their applications.

Estimating CATE

Accurately estimating CATE is fundamental to various applications. Early statistical methods were ranged from LASSO-based estimators [23] to causal forests [24]. Subsequent research extended these ideas, developing methods for censored data [25] and anomaly detection [26].

A key development was the formalization of meta-learners, flexible estimation strategies like T-learners, S-learners and X-learners [27]. More recently, neural networks have emerged as a powerful framework for CATE estimation, leading to numerous specialized architectures [28–30].

Recent work has extended CATE estimation to transformer-based architectures, leveraging attention mechanisms to model complex dependencies [31–33]. While Nadaraya-Watson kernel regression provides a theoretically grounded approach to CATE estimation [34, 35], its practical application is often limited by data sparsity, particularly in the treatment group.

Background

Concept-based learning

The paradigm of CBL formalizes a ML problem where a model must reason using a set of understandable, high-level concepts in addition to, or instead of, raw input features [36]. Formally, this framework assumes the availability of a training dataset $\mathcal{D}=\left\{\left(\mathbf{x}_i,\mathbf{c}_i,y_i\right)\right\}_{i=1}^{N}$, where $\mathbf{x}_i \in \mathcal{X} \subset \mathbb{R}^d$ is a d-dimensional input feature vector; $y_i \in \mathcal{Y} \subset \mathbb{R}$ is a continuous-valued regression target; $\mathbf{c}_i = \left(c_i^{(1)},\ldots,c_i^{(m)}\right) \in \mathcal{X} \subset \mathbb{R}^d$, in particular $\mathbf{c}_i \in \{0, 1\}^m$ is a vector of m binary concept annotations associated with $\mathbf{x}_i$. Here, $c_i^{(j)}=1$ indicates the presence of the j-th concept in the sample.

The core objective of CBL is twofold. The primary goal is to learn a hypothesis $h:\mathcal{X}\to\left(\mathcal{C},\mathcal{Y}\right)$ that can accurately predict both the target variable and the underlying concepts for a new input. The secondary, and equally critical, goal is to achieve a high degree of model explainability. By leveraging concepts as intermediate representations, CBL provides a transparent interface through which a user can understand *which concepts* present in an input were most influential in arriving at a final prediction, and to what degree they influence the predicted value.

A seminal architecture that instantiates this paradigm is the CBM [4]. The CBM explicitly decomposes the function h into two distinct stages: a concept encoder $g:\mathcal{X}\to\mathcal{C}$ that maps the raw input $\mathbf{x}$ to a vector of predicted concepts $\hat{\mathbf{c}}$; a *label predictor* $f:\mathcal{C}\to\mathcal{Y}$ that maps the predicted concepts to a final, continuous target $\hat{y}$.

The final prediction for a new input $\mathbf{x}$ is thus computed as $\hat{y}=f\left(g\left(\mathbf{x}\right)\right)$. This architectural design imposes a "concept bottleneck": all information from the input must flow through the intermediate concept representation before a final prediction is made. This ensures that the model's output is intrinsically and interpretably linked to the human-defined concepts, allowing a user to trace the predicted value $\hat{y}$ back to the specific concepts $\hat{\mathbf{c}}$ that caused it, thereby fulfilling the central promise of CBL to provide explainable predictions.

Treatment effect estimation

Let the available data be partitioned into two groups: a control group and a treatment group. The control group consists of c patients and is denoted as $\mathcal{C}=\left\{\left(\mathbf{x}_1,y_1\right),\ldots,\left(\mathbf{x}_c,y_c\right)\right\}$, where each patient i is characterized by a M-dimensional feature vector $\mathbf{x}_i=\left(x_{i1},\ldots,x_{id}\right)\in\mathbb{R}^M$ and a continuous outcome $y_i \in \mathbb{R}$ (e.g., survival time, blood pressure). Similarly, the treatment group contains t patients and is denoted as $\mathcal{T}=\left\{\left(\mathbf{z}_1,h_1\right),\ldots,\left(\mathbf{z}_t,h_t\right)\right\}$, with feature vectors $\mathbf{z}_i \in \mathbb{R}^M$ and outcomes $h_i \in \mathbb{R}$. For notational consistency across all $n = c + t$ patients, we define the treatment assignment indicator $T_i \in \{0, 1\}$, where $T_i = 0$ indicates assignment to the control group and $T_i = 1$ to the treatment group.

The central goal of causal inference is to estimate the effect of a treatment on an outcome. For a given patient, we define two *potential outcomes*: Y (the outcome if the patient does not receive the treatment, $T = 0$) and H (the outcome if the patient does receive the treatment, $T = 1$). A fundamental problem in causal inference is that for any single patient, we can only observe one of these potential outcomes, either Y or H, but never both. To overcome this, we estimate the *CATE*, which is the expected treatment effect for a subpopulation defined by a specific feature vector $\mathbf{x}$ [37]:

$$\tau\left(\mathbf{x}\right)=\mathbb{E}\left[H-Y\middle|\mathbf{X}=\mathbf{x}\right].$$

One of the important concepts if the treatment effect is the *propensity score* $e(\mathbf{x})$ which is the probability that a specific patient will receive the treatment given its observed characteristics (covariates), i.e., there holds $e(\mathbf{x}) = \Pr(T = 1 | \mathbf{X} = \mathbf{x})$. It is used to adjust differences between treatment and control groups and to isolate the true effect of a treatment from the effects of pre-existing differences.

Under some assumptions [37], CATE can be identified from the observed data as the difference between two conditional expectations:

$$\tau(\mathbf{x}) = \mathbb{E}\left[H | T = 1,\ \mathbf{X} = \mathbf{x} \right] - \mathbb{E}\left[Y | T = 0,\ \mathbf{X} = \mathbf{x} \right].$$

Let the outcome for a control patient be governed by g_0: $\mathbb{R}^d \to \mathbb{R}$ and for a treated patient by g_1: $\mathbb{R}^d \to \mathbb{R}$. Then we can write

$$y = g_0(\mathbf{x}) + \varepsilon,\ \ \mathbf{x} \in \mathcal{C},\ \ h = g_1(\mathbf{z}) + \varepsilon,\ \ \mathbf{z} \in \mathcal{T},$$

where ε is a random noise variable with $\mathbb{E}[\varepsilon] = 0$.

Hence, the CATE is simply the difference between these two response surfaces:

$$\tau(\mathbf{x}) = g_1(\mathbf{x}) - g_0(\mathbf{x}).$$

Proposed model

The proposed model, CATE-CBM, can be regarded as a combination of a CATE estimation model and a CBM. Its architecture is inspired by the Dragonnet model [38], which was introduced for CATE estimation. The architecture of CATE-CBM is depicted in Fig. 1.

It can be seen from the figure that the convolutional neural network (CNN) extracts a feature vector $\mathbf{v}$ which is fed to fully-connected neural networks (FCN-0 and FCN-1) for predicting the concept probability distributions $\mathbf{p} = (p_1, \ldots, p_m)$ for controls and $\mathbf{q} = (q_1, \ldots, q_m)$ for treatments, respectively. The use of CNNs is important for reducing the dimensionality of images. The whole network has three heads: two heads predict targets y and h from the corresponding concept probabilities; the third head can be regarded as the propensity score regularization. It forces the model to learn the structure of the confounding [38]. We propose to implement the propensity score regularization by means of the attention mechanism. In this case, the propensity score $e(\mathbf{x})$ or $e(\mathbf{z})$ is computed through the attention weights $a(\mathbf{q}_i, \mathbf{q}_j, \theta)$ with trainable parameters θ as follows:

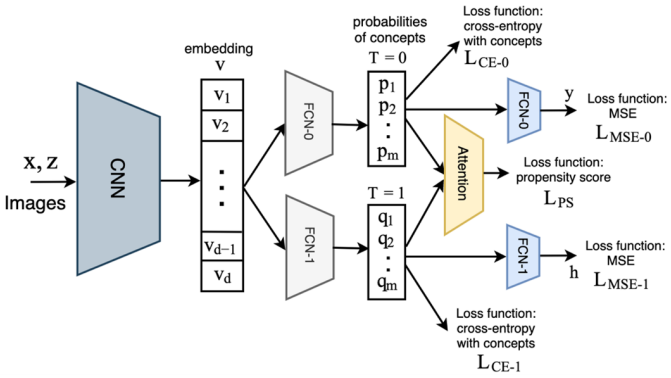


Fig. 1. Architecture of the proposed model CATE-CBM

$$e(\mathbf{x}_i) = \sum_{j=1, j\neq i}^{n} a(\mathbf{q}_i, \mathbf{q}_j, \theta) \cdot T_j,$$

where the attention weights are defined as:

$$a(\mathbf{q}_i, \mathbf{q}_j, \theta) = \frac{K_\theta(\mathbf{q}_i, \mathbf{q}_j)}{\sum_{k=1, j\neq i}^{n} K_\theta(\mathbf{q}_i, \mathbf{q}_k)}.$$

Here $K_\theta(\mathbf{q}_i, \mathbf{q}_j)$ is a kernel. In particular, if the kernel is Gaussian, then the attention weight can be expressed through the softmax function as:

$$a(\mathbf{q}_i, \mathbf{q}_j, \theta) = softmax(-\mathbf{q}_i - \mathbf{q}_j^2/\theta).$$

It should be noted that instead of the Gaussian kernel, we can use a neural network to learn a complex, data-driven similarity metric. However, this replacement significantly complicates the propensity score regularization.

In the context of the attention mechanism [39], the vector $\mathbf{q}_i$ is referred to as the *query*, while vectors $\mathbf{q}_j$ and indicators T_j are called the *keys* and *values*, respectively.

Predictions y and h for every $\mathbf{x}$ and $\mathbf{z}$, respectively, are obtained as outputs of the corresponding FCNs.

The loss function for training the whole model consists of the following five components:

• $\mathcal{L}_{MSE-0}$ and $\mathcal{L}_{MSE-1}$ are the Mean Squared Error (MSE) loss functions for the outputs y and h, respectively. The loss functions are of the form:

$$\mathcal{L}_{MSE-0} = \frac{1}{c}\sum_{i=1}^{c}(y_i - \hat{y}_i), \quad \mathcal{L}_{MSE-1} = \frac{1}{t}\sum_{i=1}^{t}(h_i - \hat{h}_i),$$

where $\hat{y}_i$ and $\hat{h}_i$ are predicted values of y_i and h_i, respectively.

• $\mathcal{L}_{CE-0}$ and $\mathcal{L}_{CE-1}$ are the cross-entropy functions controlling probabilities of concepts $\mathbf{p}$ and $\mathbf{q}$, respectively. The loss functions correspond to solving the concept classification task. For a single example with true concept values $(c_1, \ldots, c_m)$, the loss $\mathcal{L}_{CE-0}$ is defined as:

$$\mathcal{L}_{CE-0} = -\frac{1}{c}\sum_{i=1}^{c}\log(\hat{p}_i).$$

Here $\hat{p}_i$ is the predicted value of p_i. The loss $\mathcal{L}_{CE-1}$ is defined in the same way replacing $\hat{p}_i$ with $\hat{q}_i$.

• $\mathcal{L}_{PS}$ is the binary cross-entropy loss for the propensity score:

$$\mathcal{L}_{PS} = -\sum_{i=1}^{n}\left(T_i \cdot \log(\hat{e}(\mathbf{z}_i)) + (1 - T_i) \cdot \log(\hat{e}(\mathbf{x}_i))\right),$$

where $\hat{e}$ is the predicted propensity score value.

In sum, the whole loss function is defined as

$$\mathcal{L} = \gamma_1\mathcal{L}_{MSE-0} + \gamma_2\mathcal{L}_{MSE-1} + \gamma_3\mathcal{L}_{CE-0} + \gamma_4\mathcal{L}_{CE-1} + \gamma_5\mathcal{L}_{PS},$$

where γ_i, $i = 1,\ldots,5$, are hyperparameters weighting the loss components.

The predicted value of the treatment effect is computed as $\hat{\tau}(\mathbf{x}) = \hat{h}(\mathbf{x}) - \hat{y}(\mathbf{x})$. The model is trained in the end-to-end manner.

During inference, an image $\mathbf{x}$ is fed into a CNN, the output of which is an embedding $\mathbf{v}$. This embedding is then passed to two FCNs to obtain the concept probability vectors $\mathbf{p}$ and $\mathbf{q}$, such that one vector corresponds to the control group and the other to the treatment group. These vectors are fed into neural networks that generate the predictions of y and h.

It is important to note that an explicit embedding step is not strictly necessary. The concept probability vectors $\mathbf{p}$ and $\mathbf{q}$ can be obtained directly as the output of the CNN, bypassing the intermediate embedding representation. However, the intermediate embedding may have advantage in comparison with the direct implementation of the concept probabilities as the output of the CNN. The embedding layer can act as a form of regularization, preventing the concept predictors from overfitting to the training data by forcing information compression. Moreover, the CNN's feature maps are often low-level (edges, textures). An embedding layer can learn to combine these low-level features into a more sophisticated, high-level representation that is better suited for predicting complex concepts.

Local interpretation of the CATE predictions

An important question in CATE estimation is its interpretation that lies in answering the question of which concept change had the strongest impact on the estimated CATE value. The proposed model allows us to answer this question in the following way.

First, it should be noted that we consider the local interpretation which allows us to explain an individual prediction at a point of interest. Methods of the local interpretation are based on a linear approximation of the predictive model in a neighborhood around the explainable point [40, 41]. A well-known local explanation method is the Local Interpretable Model-agnostic Explanations (LIME) [42] interpreting the black-box model predictions by approximating the model at a point by a linear model whose coefficients can be viewed as a quantitative representation of the feature impacts on the prediction [43]. The approach to interprete predictions by means of the black-box model approximation at a point by the linear model can be applied to many classification and regression tasks. Therefore, we consider its use for interpreting the CATE predictions.

The output FCN-0 and FCN-1 are linear, which makes it possible to interpret each predicted function $\hat{y}(\mathbf{x})$ and $\hat{h}(\mathbf{x})$ in terms of concepts. If we identify the concepts with the highest probability, then the largest weights of the neural networks precisely show which of the identified concepts are significant for each function. It should be noted that the absence of a concept can also be significant. The change in their values, if present for a given concept, is exactly what answers the question of CATE interpretation.

Suppose that functions $f_0(\mathbf{x}) = \mathbf{a}^{\mathrm{T}}\mathbf{x} + a_0$ and $f_1(\mathbf{x}) = \mathbf{b}^{\mathrm{T}}\mathbf{x} + b_0$ are linear approximations of functions $\hat{y}(\mathbf{x})$ and $\hat{h}(\mathbf{x})$, respectively, at a point $\mathbf{x}$, where $\mathbf{a} = (a_1,...,a_d)^{\mathrm{T}}$ and $\mathbf{b} = (b_1,...,b_d)^{\mathrm{T}}$. Then we can write $\hat{y}(\mathbf{x}) = f_0(\mathbf{x}) + e_0$ and $\hat{h}(\mathbf{x}) = f_1(\mathbf{x}) + e_1$. Here e_0 and e_1 are the approximation errors. The CATE in this case is defined as

$$\hat{\tau}(\mathbf{x}) = \hat{h}(\mathbf{x}) - \hat{y}(\mathbf{x}) = f_1(\mathbf{x}) - f_0(\mathbf{x}) + e_1 - e_2.$$

Note that there holds

$$|e_1 - e_0| \le |e_1| + |e_0|.$$

Hence, the linear difference $f_1(\mathbf{x}) - f_0(\mathbf{x})$ can be regarded as a linear approximation of $\hat{\tau}(\mathbf{x})$ at point $\mathbf{x}$ if the approximation errors e_1 and e_0 are small. Finally, if we know weights of FCN-0 and FCN-1, $\mathbf{a}$ and $\mathbf{b}$, the the values of the concept importance are defined as the difference $\mathbf{b} - \mathbf{a}$.

Numerical experiments

A difficulty of comparing the proposed model with other models is that CATE-CBM is the first model combining CATE and CBM. Therefore, we will show some properties of CATE-CBM by means of numerical experiments.

To study the proposed model, a synthetic dataset is constructed from the well-known MNIST dataset [44] which represents 28×28 pixel handwritten digit images. The original MNIST dataset has a training set of 60000 instances and a test set of 10000 instances[1].

Each instance in the synthetic dataset consists of four different digits randomly taken from MNIST such that the instance has two digits in the first row and two digits in the second row as it is shown in Fig. 2.

Each instance has the size 56×56. A similar dataset is used in [14] and in [45]. Concepts $c^{(1)},...,c^{(10)}$ are binary and defined by the presence of the corresponding number 1,...,9,0 in the instance. For example, the first instance has concepts (1,1,1,0,0,0,0,0,0,1), the second instance has concepts (0,0,0,0,1,1,1,0,0,1).

We analyze the proposed model by its training on numbers of instances (from 1000 till 5000). The number of testing images is 20000. The cross-validation in all experiments is performed with 50 repetitions.

For experiments, we apply functions similar to those used in [27]. They are expressed through the indicator function I taking value 1 if its argument is true. The function for controls is represented as

$$g_0(\mathbf{c}) = \mathbf{b}^{\mathrm{T}}\mathbf{c} + 5I(c_4 = 1),$$

where $\mathbf{b}^{\mathrm{T}} = (1,2,3,4,5,4,3,2,1,0.01)$.

The function for treatments is represented as

$$g_1(\mathbf{c}) = \mathbf{b}^{\mathrm{T}}\mathbf{c} + 5I(c_3 = 1) + 7I(c_8 = 1),$$

where $\mathbf{b}^{\mathrm{T}} = (0.01,1,2,3,4,5,4,3,2,1)$.

Values of y and h are generated by adding the normally distributed random numbers ε with the zero expectation and the standard deviations $\sigma_0 = 1.5$ for controls and $\sigma_1 = 2.0$ for treatments.

Suppose that we have additional information about peculiarities of digits in controls and treatments, namaly controls do not have digits “1”, treatments do not have digits “7”. In accordance with this information, we generate examples for control and treatment groups. It can be seen from Fig. 2 that the first example belongs to treatments, the second belongs to controls, the third and fouth can belong to treatments as well as to controls, therefore, they are randomly referred to the controls or treatments. If an example contains “1” and “7” simultaneously, then it is removed from the generated dataset.

For the modified MNIST datasets, we employ CNN (Fig. 1) consisting of four convolutional layers with progressively decreasing kernel sizes, starting from (8×8) to (4×4). LeakyReLU activation functions are used throughout, and the final layer is linear. FCN-0 and FCN-1 consist of two layers with sigmoid activation functions. The output FCNs contain one layer which is linear to implement interpretability of the CATE. Parameter θ of the Gaussian kernel in the propensity score regularization is trainable.

The MSE measure is used as an accuracy measure of CATE in experiments. We compare the proposed model CATE-CBM with the same implementation of the CATE estimator but without using concepts. MSE of CATE with and without using concepts is shown in Fig. 3, where MSE as functions

[1] The dataset is available at http://yann.lecun.com/exdb/mnist/.

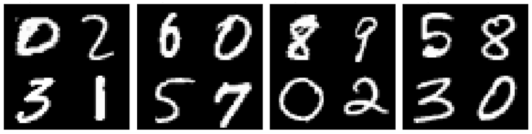


Fig. 2. Examples of the modified MNIST dataset

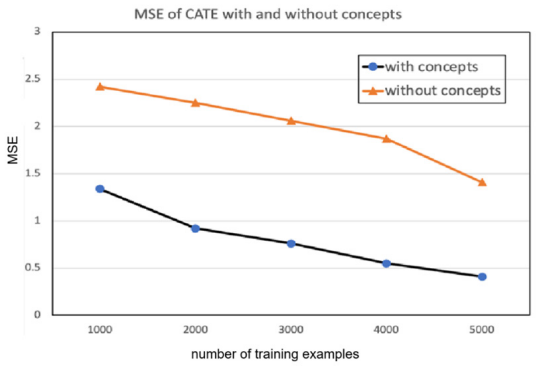


Fig. 3. MSE of CATE with and without using concepts

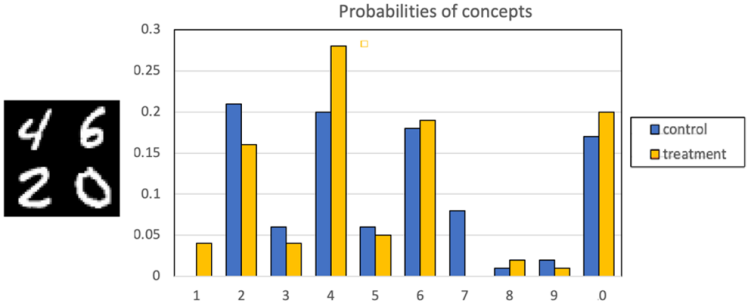


Fig. 4. Probabilities of concepts for the testing example

of the training set size of the modified MNIST is demonstrated. It can be seen from the graphs that information about concepts significantly improves the model performance.

Another important consideration is the interpretation of the results. Following the aforementioned interpretation method, we compute the probabilities of concepts under the assumptions that an example belongs to the control and treatment groups. Fig. 4 illustrates these concept probabilities for an example containing the digits 4, 6, 2 and 0. The probabilities for these corresponding concepts are the largest, indicating that the CATE-CBM model correctly recognizes them.

Fig. 5 shows the normalized weights from the output layers FCN-0 and FCN-1. The importance of each concept for the CATE prediction can be derived from the difference between the weights of FCN-1 and FCN-0. The results, depicted in Fig. 6, reveal that the most important concept is “5”, which is not present in the example. This interesting finding demonstrates that the absence of a concept can also be significant. Furthermore, the importance value for this concept is negative, implying that concept 5 acts to reduce the treatment effect. In contrast, concepts 3 and 7 have positive importance.

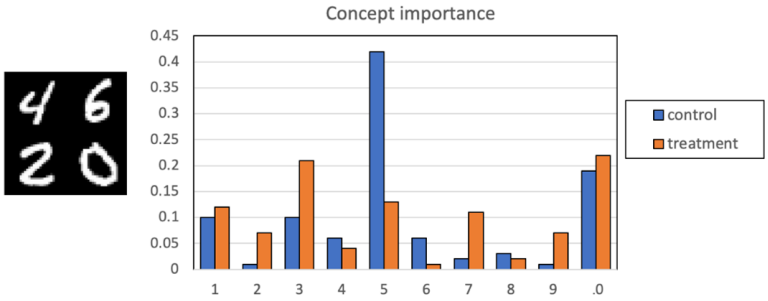


Fig. 5. Importance of concepts for the testing example

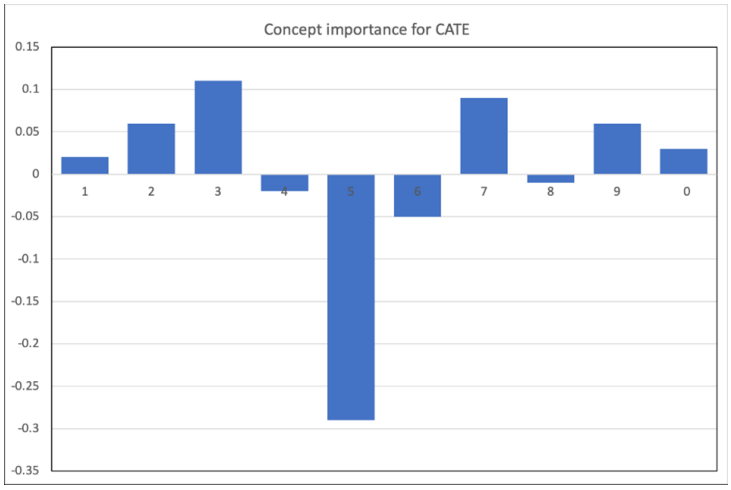


Fig. 6. Values of the concept importance for the CATE interpretation of the considered example

It should be noted that the above interpretation results pertain only to the specific example with digits 4, 6, 2 and 0.

Conclusion

This paper introduced CATE-CBM, the first model to integrate CBL with CATE estimation. By combining interpretable concept bottlenecks with CATE estimation, the model provides both accurate treatment effect predictions and understandable explanations through concept importance analysis.

The numerical experiments conducted on the modified MNIST dataset demonstrate several important properties and advantages of the proposed CATE-CBM model.

1. First, the incorporation of concept information significantly enhances model accuracy, as evidenced by the consistently lower MSE of CATE estimation compared to the same model without concept utilization.

2. Second, the model successfully identifies and extracts relevant concepts from complex image data, as shown by the high probability scores assigned to correct digit concepts in test examples.

3. Third, CATE-CBM provides transparent insights into treatment effect mechanisms through concept importance analysis.

4. Fourth, the CNN-Concept architecture proves effective for handling complex visual data while maintaining interpretability, successfully balancing predictive performance with explanatory capabilities.

These findings establish CATE-CBM as a promising approach for CATE estimation in settings where both accuracy and interpretability are crucial, particularly when dealing with high-dimensional data requiring meaningful feature extraction.

Several promising directions emerge from this work.

1. First, extending CATE-CBM to handle continuous-valued concepts and temporal treatment effects would broaden its applicability.

2. Second, developing methods for automatic concept discovery rather than relying on pre-defined concepts could enhance model flexibility.

3. Third, incorporating uncertainty quantification for both concept predictions and treatment effects would provide crucial reliability measures for decision-making.

Applications in real-world clinical trials and policy evaluation settings would further validate the approach's practical utility.

REFERENCES

1. **Kim B., Wattenberg M., Gilmer J, Cai C., Wexler J., Viegas F., Sayres R.** Interpretability beyond feature attribution: Quantitative Testing with Concept Activation Vectors (TCAV). *Proceedings of the 35th International Conference on Machine Learning*, 2018, Vol. 80, Pp. 2668–2677. DOI: 10.48550/arXiv.1711.11279

2. **Yeh C.-K., Kim B., Arik S.Ö., Li C.-L., Pfister T., Ravikumar P.** On completeness-aware concept-based explanations in deep neural networks. *Proceedings of the 34th International Conference on Neural Information Processing Systems* (*NeurIPS 2020*), 2020, Pp. 20554–20565.

3. **Wang B., Li L., Nakashima Y., Nagahara H.** Learning bottleneck concepts in image classification. *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*), 2023, Pp. 10962–10971. DOI: 10.1109/CVPR52729.2023.01055

4. **Koh P.W., Nguyen T., Tang Y.S., Mussmann S., Pierson E., Kim B., Liang P.** Concept bottleneck models. *Proceedings of the 37th International Conference on Machine Learning*, 2020, Vol. 119, Pp. 5338–5348.

5. **Caron A., Baio G., Manolopoulou I.** Estimating individual treatment effects using non-parametric regression models: A review. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 2022, Vol. 185, No. 3, Pp. 1115–1149. DOI: 10.1111/rssa.12824

6. **Künzel S.R., Stadie B.C., Vemuri N., Ramakrishnan V., Sekhon J.S., Abbeel P.** Transfer learning for estimating causal effects using neural networks. *arXiv:1808.07804*, 2018. DOI: 10.48550/arXiv.1808.07804

7. **Zhou X., Xie Y.** Heterogeneous treatment effects in the presence of self-selection: A propensity score perspective. *Sociological Methodology*, 2020, Vol. 50, No. 1, Pp. 350–385. DOI: 10.1177/0081175019862593

8. **Chu Z., Li S.** Continual treatment effect estimation: Challenges and opportunities. *Proceedings of Machine Learning Research*, 2023, Vol. 208, Pp. 11–17.

9. **Curth A., Peck R.W., McKinney E., Weatherall J., van der Schaar M.** Using machine learning to individualize treatment effect estimation: Challenges and opportunities. *Clinical Pharmacology & Therapeutics*, 2024, Vol. 115, No. 4, Pp. 710–719. DOI: 10.1002/cpt.3159

10. **Wang Y., Li H., Zhu M., Wu A., Xiong R., Wu F., Kuang K.** Causal inference with complex treatments: A survey. *arXiv:2407.14022*, 2024. DOI: 10.48550/arXiv.2407.14022

11. **Yao L., Chu Z., Li S., Li Y., Gao J., Zhang A.** A survey on causal inference. *ACM Transactions on Knowledge Discovery from Data* (*TKDD*), 2021, Vol. 15, No. 5, Art. no. 74. DOI: 10.1145/344494

12. **Zhang W., Li J., Liu L.** A unified survey of treatment effect heterogeneity modelling and uplift modelling. *ACM Computing Surveys* (*CSUR*), 2021, Vol. 54, No. 8, Art. no. 162. DOI: 10.1145/3466818

13. **Ismail A.A., Adebayo J., Bravo H.C., Ra S., Cho K.** Concept bottleneck generative models. *Proceedings of ICML 2023*. Workshop on Deployment Challenges for Generative AI, pages 1–10, 2023.

14. **Kim E., Jung D., Park S., Kim S., Yoon S.** Probabilistic concept bottleneck models. *The 12th International Conference on Learning Representations*, 2024.

15. **Raman N., Zarlenga M.E., Heo J., Jamnik M.** Do concept bottleneck models obey locality? *XAI in Action: Past, Present, and Future Applications*, 2023.

16. **Laguna S., Marcinkevičs R., Vandenhirtz M., Vogt J.E.** Beyond concept bottleneck models: How to make black boxes intervenable? *arXiv:2401.13544*, 2024. DOI: 10.48550/arXiv.2401.13544

17. **Kazmierczak R., Berthier E., Frehse G., Franchi G.** CLIP-QDA: An explainable concept bottleneck model. *Transactions on Machine Learning Research Journal*, 2024.

18. **Radford A., Kim J.W., Hallacy C., Ramesh A., Goh G., Agarwal S., Sastry G., Askell A., Mishkin P., Clark J., Krueger G., Sutskever I.** Learning transferable visual models from natural language supervision. *Proceedings of the 38th International Conference on Machine Learning*, 2021, Vol. 139, Pp. 8748–8763.

19. **Havasi M., Parbhoo S., Doshi-Velez F.** Addressing leakage in concept bottleneck models. *36th Conference on Neural Information Processing Systems* (*NeurIPS*), 2022.

20. **Gupta A., Narayanan P.J.** A survey on concept-based approaches for model improvement. *arXiv:2403.14566*, 2024. DOI: 10.48550/arXiv.2403.14566

21. **Lee J.H., Lanza S., Wermter S.** From neural activations to concepts: A survey on explaining concepts in neural networks. *arXiv:2310.11884*, 2023. DOI: 10.48550/arXiv.2310.11884

22. **Mahinpei A., Clark J., Lage I., Doshi-Velez F., Pan W.** Promises and pitfalls of black-box concept learning models. *arXiv:2106.13314*, 2021. DOI: 10.48550/arXiv.2106.13314

23. **Jeng X.J., Lu W., Peng H.** High-dimensional inference for personalized treatment decision. *Electronic Journal of Statistics*, 2018, Vol. 12, No. 1, Pp. 2074–2089. DOI: 10.1214/18-EJS1439

24. **Athey S., Tibshirani J., Wager S.** Generalized random forests. *arXiv:1610.0171*, 2016. DOI: 10.48550/arXiv.1610.01271

25. **Zhang W., Le T.D., Liu L., Zhou Z.-H., Li J.** Mining heterogeneous causal effects for personalized cancer treatment. *Bioinformatics*, 2017, Vol. 33, No. 15, Pp. 2372–2378. DOI: 10.1093/bioinformatics/btx174

26. **McFowland III E., Somanchi S., Neill D.B.** Efficient discovery of heterogeneous treatment effects in randomized experiments via anomalous pattern detection. *arXiv:1803.09159*, 2018. DOI: 10.48550/arXiv.1803.09159

27. **Künzel S.R., Sekhon J.S., Bickel P.J., Yu B.** Meta-learners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the National Academy of Sciences U.S.A.*, 2019, Vol. 116, No. 10, Pp. 4156–4165. DOI: 10.1073/pnas.1804597116

28. **Curth A., van der Schaar M.** Nonparametric estimation of heterogeneous treatment effects: From theory to learning algorithms. *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics* (*AISTATS*), 2021, Vol. 130, Pp. 1–11.

29. **Du X., Fan Y., Lv J., Sun T., Vossler P.** Dimension-free average treatment effect inference with deep neural networks. *arXiv:2112.01574*, 2021. DOI: 10.48550/arXiv.2112.01574

30. **Qin T., Wang T.-Z., Zhou Z.-H.** Budgeted heterogeneous treatment effect estimation. *Proceedings of the 38th International Conference on Machine Learning*, 2021, Vol. 139, Pp. 8693–8702.

31. **Guo Z., Zheng S., Liu Z., Yan K., Zhu Z.** CETransformer: Casual effect estimation via transformer based representation learning. *Pattern Recognition and Computer Vision: 4th Chinese Conference*, 2021, Vol. IV, Pp. 524–535. DOI: 10.1007/978-3-030-88013-2_43

32. **Melnychuk V., Frauen D., Feuerriegel S.** Causal transformer for estimating counterfactual outcomes. *arXiv:2204.07258*, 2022. DOI: 10.48550/arXiv.2204.07258

33. **Zhang Y.-F., Zhang H., Lipton Z.C., Li L., Xing E.P.** Exploring transformer backbones for heterogeneous treatment effect estimation. *arXiv:2202.01336*, 2022. DOI: 10.48550/arXiv.2202.01336

34. **Imbens G.W.** Nonparametric estimation of average treatment effects under exogeneity: A review. *The Review of Economics and Statistics*, 2004, Vol. 86, No. 1, Pp. 4–29.

35. **Park J., Shalit U., Scholkopf B., Muandet K.** Conditional distributional treatment effect with kernel conditional mean embeddings and U-statistic regression. *Proceedings of the 38th International Conference on Machine Learning*, 2021, Vol. 139, Pp. 8401–8412.

36. **Xu K., Fukuchi K., Akimoto Y., Sakuma J.** Statistically significant concept-based explanation of image classifiers via model knockoffs. *Proceedings of the 32nd International Joint Conference on Artificial Intelligence*, 2023, Pp. 519–526. DOI: 10.24963/ijcai.2023/58

37. **Rubin D.B.** Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association*, Vol. 100, No. 469, Pp. 322–331. DOI: 10.1198/016214504000001880

38. **Shi C., Blei D.M., Veitch V.** Adapting neural networks for the estimation of treatment effects. *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, 2019, Pp. 2507–2517.

39. **Bahdanau D., Cho K., Bengio Y.** Neural machine translation by jointly learning to align and translate. *arXiv:1409.0473*, 2014. DOI: 10.48550/arXiv.1409.0473

40. **Bodria F., Giannotti F., Guidotti R., Naretto F., Pedreschi D., Rinzivillo S.** Benchmarking and survey of explanation methods for black box models. *Data Mining and Knowledge Discovery*, 2023, Vol. 37, Pp. 1719–1778. DOI: 10.1007/s10618-023-00933-9

41. **Burkart N., Huber M.F.** A survey on the explainability of supervised machine learning. *Journal of Artificial Intelligence Research*, 2021, Vol. 70, Pp. 245–317. DOI: 10.1613/jair.1.12228

42. **Ribeiro M.T., Singh S., Guestrin C.** "Why should I trust you?": Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, Pp. 1135–1144. DOI: 10.1145/2939672.2939778

43. **Garreau D., von Luxburg U.** Explaining the explainer: A first theoretical analysis of LIME. *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics* (*AISTATS*), 2020, Vol. 108, Pp. 1–9.

44. **LeCun Y., Bottou L., Bengio Y., Haffner P.** Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 1998, Vol. 86, No. 11, Pp. 2278–2324. DOI: 10.1109/5.726791

45. **Kirpichenko S.R., Utkin L.V., Konstantinov A.V., Verbova N.M.** Survival concept-based learning models. *Journal of Intelligent Information Systems*, 2025, Vol. 63, Pp. 1687–1711. DOI: 10.1007/s10844-025-00958-0

INFORMATION ABOUT AUTHORS / СВЕДЕНИЯ ОБ АВТОРАХ

Lev V. Utkin
Уткин Лев Владимирович
E-mail: lev.utkin@gmail.com
ORCID: https://orcid.org/0000-0002-5637-1420

Andrei V. Konstantinov
Константинов Андрей Владимирович
E-mail: andrue.konst@gmail.com
ORCID: https://orcid.org/0000-0002-1542-6480

Natalia M. Verbova
Вербова Наталья Михайловна
E-mail: nag00@mail.ru
ORCID: https://orcid.org/0000-0001-8749-9470

Submitted: 18.11.2025; Approved: 13.12.2025; Accepted: 28.12.2025.
Поступила: 18.11.2025; Одобрена: 13.12.2025; Принята: 28.12.2025.

Computing, Telecommunication and Control, 2025, Vol. 18, No. 4, Pp. 20–29.
Информатика, телекоммуникации и управление. 2025. Том 18, № 4. С. 20–29.

Research article
DOI: https://doi.org/10.18721/JCSTCS.18402
UDC 004.032.26

A STUDY OF THE APPLICABILITY OF THE KOLMOGOROV-ARNOLD NETWORK ARCHITECTURE FOR TIME SERIES FORECASTING

O.G. Maleev ✉ , *O.A. Kovaleva*

Peter the Great St. Petersburg Polytechnic University,
St. Petersburg, Russian Federation

✉ olegmg@bk.ru

Abstract. The recently proposed Kolmogorov–Arnold Network (KAN) architecture emerges as a promising alternative to traditional neural networks based on the Multilayer Perceptron (MLP). By leveraging the Kolmogorov–Arnold representation theorem, KAN represents multidimensional functions as combinations of univariate functions, thereby offering potentially higher accuracy and model interpretability through its inherently simpler structure. This paper investigates the applicability of KAN to time series forecasting using the well-known hourly electricity consumption dataset as a benchmark. Meteorological observation data are selected as an additional testbed. A comparative analysis is conducted between KAN networks and traditional MLPs, as well as implementations of recurrent architectures based on KAN (TKAN variants) versus established designs such as Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU). Experimental results demonstrate the superiority of the KAN architecture over MLPs in temporal prediction tasks. The proposed recurrent architecture, TKAN1, achieves the highest coefficient of determination ($R^2 = 0.3483$) among TKAN variants, with a Root Mean Squared Error (RMSE) of 0.1010 in energy demand forecasting.

Keywords: time series, time series forecasting, Kolmogorov–Arnold network, multilayer perceptron, recurrent neural network

Citation: Maleev O.G., Kovaleva O.A. A study of the applicability of the Kolmogorov–Arnold network architecture for time series forecasting. Computing, Telecommunications and Control, 2025, Vol. 18, No. 4, Pp. 20–29. DOI: 10.18721/JCSTCS.18402

Научная статья
DOI: https://doi.org/10.18721/JCSTCS.18402
УДК 004.032.26

ИССЛЕДОВАНИЕ ПРИМЕНИМОСТИ АРХИТЕКТУРЫ НЕЙРОННЫХ СЕТЕЙ КОЛМОГОРОВА-АРНОЛЬДА (KAN) К ЗАДАЧЕ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ

О.Г. Малеев ✉ , *О.А. Ковалева*

Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Российская Федерация

✉ olegmg@bk.ru

Аннотация. Недавно предложенная архитектура нейросетей Колмогорова–Арнольда (Kolmogorov–Arnold Networks, KAN) является перспективной альтернативой традиционным нейронным сетям на основе многослойного персептрона (Multilayer Perceptron, MLP). Благодаря использованию теоремы Колмогорова–Арнольда, KAN представляет многомерные функции в виде комбинации одномерных, обеспечивая потенциально более высокую точность и интерпретируемость модели. В данной статье исследуется применимость KAN и ее рекуррентного расширения – Temporal Kolmogorov–Arnold Networks (TKAN) – к задаче прогнозирования временных рядов на примере известного набора данных почасового потребления электроэнергии. В качестве дополнительного набора данных выбраны данные метеорологических наблюдений. Проведен сравнительный анализ сетей KAN с традиционными MLP, а также реализации архитектуры рекуррентной нейросети на основе архитектуры KAN с широко известными архитектурами долгой краткосрочной памяти (Long Short-Term Memory, LSTM) и управляемого рекуррентного блока (Gated Recurrent Units, GRU). Экспериментальные результаты демонстрируют превосходство архитектуры KAN над MLP в задачах временного прогнозирования. Предложенная в статье рекуррентная архитектура TKAN1 демонстрирует лучший среди TKAN коэффициент детерминации $R^2 = 0{,}3483$ при RMSE 0,1010 в задаче прогнозирования энергопотребления.

Ключевые слова: временные ряды, прогнозирование временных рядов, нейросеть Колмогорова–Арнольда, многослойный персептрон, рекуррентная нейронная сеть

Для цитирования: Maleev O.G., Kovaleva O.A. A study of the applicability of the Kolmogorov–Arnold network architecture for time series forecasting // Computing, Telecommunications and Control. 2025. Т. 18, № 4. С. 20–29. DOI: 10.18721/JCSTCS.18402

Introduction

Forecasting time series is a critical task across diverse industries, from economics and transportation to meteorology and medicine. Achieving highly accurate forecasts is essential for maintaining business competitiveness, minimizing risks, optimizing resources and justifying significant decision-making processes.

Established approaches address time series forecasting using classical statistical methods [1–3]. However, with increased computational power and the daily generation of vast volumes of temporal data, deep neural networks have proven effective for forecasting. These networks can learn complex data representations, eliminating or reducing the need for manual feature engineering [4, 5]. Despite notable advancements in time series forecasting, several challenges persist. Models based on the Multilayer Perceptron (MLP) architecture require larger statistical datasets for training due to the lack of prior knowledge [6]. Furthermore, black-box models exhibit reduced interpretability and explain ability compared to statistical methods.

A recently proposed fundamentally novel neural network architecture, the Kolmogorov–Arnold Network (KAN) [7], presents a promising alternative to the MLP and opens new avenues for advancing deep learning models.

Related works

MLP networks are effective approximators of nonlinear functions due to the underlying universal approximation theorem (Cybenko's theorem) [8], which asserts, that a feedforward network with a single hidden layer containing a finite number of neurons can approximate any continuous multivariate function to arbitrary accuracy, provided the hidden layer contains a sufficient number of neurons and the network parameters are appropriately chosen. The KAN, in turn, is grounded in the Kolmogorov–Arnold representation theorem, which states that any multivariate continuous function can be expressed as a composition of univariate functions and addition operations:

$$f\left(x_1,\ldots,x_n\right)=\sum_{q=1}^{2n+1}\Phi_q\left(\sum_{p=1}^{n}\phi_{q,p}\left(x_p\right)\right),$$

where $\phi_{q,p}:[0,1]\to\mathbb{R},\ \Phi_q:\mathbb{R}\to\mathbb{R}$.

Each univariate function is parameterized as a B-spline curve with trainable coefficients c_i of local B-spline basis functions $B_i\left(x\right)$, and is represented as a weighted sum with trainable weights w_b and w_s:

$$\phi\left(x\right)=w_b b\left(x\right)+w_s spline\left(x\right),\quad spline\left(x\right)=\sum_i c_i B_i\left(x\right),$$

where $b\left(x\right)$ is analogous to a residual connection.

Thus, l-th layer of KAN is defined by a matrix of functions:

$$\Phi_l=\left\{\phi_{l,q,p}\right\},\ \ l=0,\ \ldots,\ L-1,\ \ p=1,\ \ldots,\ n_{in},\ \ q=1,\ \ldots,\ n_{out},$$

enabling KAN to extend the Kolmogorov–Arnold representation theorem to arbitrary width and depth [9].

The general structure of KAN comprises a composition of L layers, with dimensions specified by the array $\left[n_0,\ \ldots,\ n_{L-1}\right]$. For an input vector $X\in\mathbb{R}^{n_0}$, the KAN output is expressed as:

$$KAN\left(X\right)=\left(\mathbf{\Phi}_{L-1}\circ\mathbf{\Phi}_{L-2}\circ\ldots\circ\mathbf{\Phi}_0\right)X,$$

where each layer $\mathbf{\Phi}_l$ transforms its input through learnable univariate functions parameterized via B-spline basis expansions.

Thus, in KAN, activation functions are moved to the edges of the computational graph: each weight is replaced by a univariate activation function parameterized as a spline, while neurons themselves perform only summation of incoming activations.

KAN combines the strengths of MLPs and splines, featuring internal and external levels of degrees of freedom. At the external level, KAN learns the compositional structure of the target function through its MLP-like architecture, while at the internal level, it approximates univariate functions with high precision via spline-like parameterization. Architectural complexity in KAN involves not only adding more layers but also refining the spline grids.

The KAN approach to representing multivariate functions aligns with structural properties of time series, such as trends and seasonality. Embedding prior knowledge about data structure directly into the neural network architecture suggests KAN's potential effectiveness in time series forecasting [10].

Transition to the recurrent TKAN architecture

To extend the capabilities of KAN for time series forecasting, a natural direction is to integrate the KAN architecture with other widely adopted deep learning methods for this task. Recurrent Neural Network (RNN) models have demonstrated themselves as effective tools for forecasting in practical scenarios of varying complexity [11].

For this study, we use a combination of a Recurrent Kolmogorov–Arnold Network (RKAN) and a modified Long Short-Term Memory (LSTM) block, forming the Temporal Kolmogorov–Arnold Network (TKAN) architecture [12]. This approach enables capturing complex nonlinearities through RKAN's learnable activation functions while efficiently managing memory over extended periods via the LSTM cell architecture.

KAN layers retain short-term memory of previous network states, and the gating mechanism regulates information flow by determining which information should be preserved or forgotten over time. The schematic of the TKAN cell is illustrated in Fig. 1.

Following the analogy of hidden state updates in RNNs, the dependence of the current hidden state on its prior value introduces temporal dynamics into each activation function $\phi_{l,j,i}$:

$$x_{l+1,j}(t) = \sum_{i=1}^{n_l} \tilde{x}_{l,j,i}(t) = \sum_{i=1}^{n_l} \phi_{l,j,i}\left(x_{l,i}(t), h_{l,i}(t)\right), \quad j = 1, \ldots, n_{l+1},$$

where $h_{l,i}(t)$ is the memory state function for the i-th neuron of the l-th layer at time t.

By analogy with the LSTM cell, information flow in the TKAN cell is governed by a forget gate, input gate and output gate:

$$f_t = \sigma\left(W_f x_t + U_t h_{t-1} + b_f\right), \; i_t = \sigma\left(W_i x_t + U_i h_{t-1} + b_i\right),$$
$$o_t = \sigma\left(RKAN\left(X, t\right)\right).$$

The hidden state h_t is computed as the output of the cell:

$$h_t = o_t \odot \tanh\left(c_t\right),$$

where c_t represents the long-term memory of the cell, updated according to:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t,$$

where $\tilde{c}_t = \sigma\left(W_c x_t + U_c h_{t-1} + b_c\right)$.

The final predicted value $\hat{y}_t$ is derived via a linear layer:

$$\hat{y}_t = W_{hy} h_t + b_y.$$

This formulation aligns with standard LSTM-based memory mechanisms, where the hidden state h_t acts as a compressed representation of temporal dependencies, and the cell state c_t retains long-term memory through gated updates.

Numerical experiments

To evaluate the applicability of KAN to time series forecasting, we conduct a comparative analysis of KAN against MLP and TKAN against classical recurrent architectures – LSTM [13] and GRU [14].

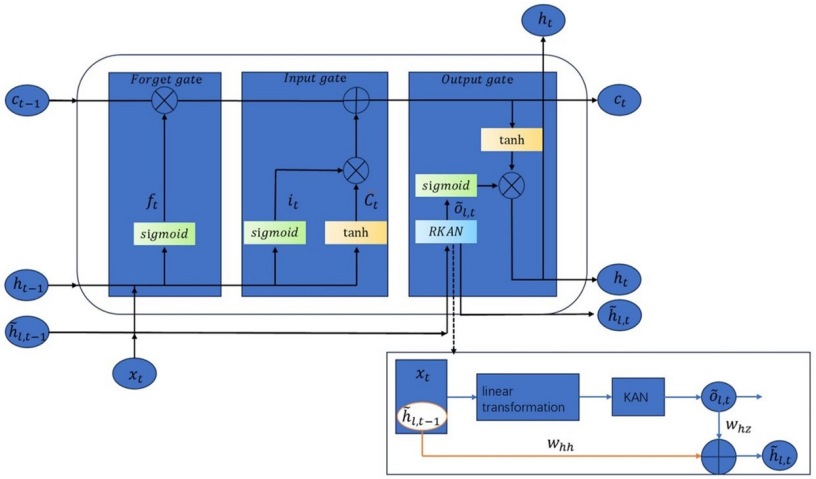


Fig. 1. Architecture of the TKAN Cell

We frame multivariate time series forecasting as a supervised learning task. The input to the model is a sequence of historical time steps $X = [X_1, \ldots, X_H] \in \mathbb{R}^{H \times D}$, where H denotes the historical window size used for forecasting and D represents the number of variables. The forecasting task involves generating an output sequence $Y = [X_{H+1}, \ldots, X_{H+F}] \in \mathbb{R}^{F \times D}$, where F is the forecasting horizon.

We evaluate the considered models on widely used benchmark datasets (Table 1):

1. Electricity Dataset: This dataset contains hourly electricity consumption data from 321 clients between 2012 and 2014, measured in kilowatts[1].

2. Weather Dataset: This dataset includes meteorological observations recorded at 10-minute intervals near Beutenberg (Germany) from 2021 to 2023[2]. Hourly data was obtained by sampling the first observation of each hour.

Table 1

Characteristics of the used datasets

	Dimensions	Series length	Granularuty	Split
Electricity	321	26304	1 hour	[7:1:2]
Weather	20	26280	1 hour	[8:0.5:1.5]

Data preprocessing involves MinMax scaling, which maps values to the [0,1] interval while preserving the distribution shape. Validation follows a simple strategy: datasets are split into training, validation and testing subsets in chronological order according to the predefined ratios.

We compare MLP and KAN on the Electricity dataset, analyzing the dependence of model performance on a key KAN parameter – the spline grid size G. Identical network configurations are considered, with varying numbers of hidden layers and neurons. The historical window size is fixed at $H = 24$, and the forecasting horizon is set to $F = 6$. Network configurations are detailed in Table 2. The loss function is Mean Squared Error (MSE), and training employs the Adam optimizer with an initial learning rate of lr = 0.001. The MSE and Mean Absolute Error (MAE) are used as evaluation metrics for models on test data. Results are summarized in Table 2.

[1] Electricity Hourly Dataset, Available: https://zenodo.org/records/4656140 (Accessed 08.04.2025)

[2] Max-Planck-Institut fuer Biogeochemie – Wetterdaten, Available: https://www.bgc-jena.mpg.de/wetter/weather_data.html (Accessed 08.04.2025)

Table 2

Comparison of MLP and KAN evaluation results

	Configuration	MSE	MAE	Number of parameters
KAN	[24, 5, 6], G=3	0.00471	0.04299	1.2 K
	[24, 10, 6], G=3	0.00451	0.04215	2.4 K
	[24, 20, 6], G=3	0.00442	0.04180	4.8 K
	[24, 10, 10, 6], G=3	0.00464	0.04354	3.2 K
	[24, 10, 6], G=10	0.00419	**0.04060**	4.5 K
	[24, 10, 10, 6], G=10	**0.00416**	0.04081	6.0 K
MLP	[24, 5, 6]	0.00561	0.04801	161
	[24, 10, 6]	0.00498	0.04484	316
	[24, 20, 6]	0.00460	0.04265	626
	[24, 10, 10, 6]	0.00522	0.04655	426
	[24, 20, 20, 20, 6]	0.00480	0.04440	1.5 K

For TKAN, LSTM and GRU, we adopt a simple unified architecture: an input recurrent layer returning full sequences and an intermediate layer returning only the final hidden state, both with identical hidden state dimensions. The output layer is fully connected with linear activation. Hidden state sizes $h = [50, 100]$, KAN layer sizes $h_k = 20$, spline grid size $G = 3$ and spline order $k = 3$ are tested. Forecasts are generated for $F = [12, 24, 48, 96, 168]$ with $H = 48$. The loss function remains MSE, and evaluation metrics include RMSE, MAE and the coefficient of determination R^2. Results are reported in Tables 3–6, with bold indicating the best performance per metric and forecasting horizon.

Table 3

Evaluation results of recurrent models with hidden state size $h = 100$ on the electricity dataset

F	TKAN			GRU			LSTM		
	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE
12	0.1316	**0.0766**	**0.0593**	0.2228	0.0797	0.0621	**0.3497**	0.0767	0.0596
24	0.2073	0.0767	0.0590	0.3592	**0.0756**	**0.0585**	**0.3668**	0.0777	0.0605
48	0.2292	0.0873	0.0677	**0.3204**	**0.0776**	**0.0597**	0.2436	0.0808	0.0624
96	**0.2869**	0.0829	0.0640	0.2790	0.0829	0.0639	0.1663	**0.0822**	**0.0632**

Table 4

Evaluation results of recurrent models with hidden state size $h = 100$ on the weather dataset

F	TKAN			GRU			LSTM		
	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE
12	0.6822	0.0738	0.0506	**0.7029**	**0.0697**	**0.0469**	0.6676	0.0747	0.0514
24	0.5893	0.0884	0.0633	0.6070	0.0858	0.0611	**0.6133**	**0.0842**	**0.0590**
48	0.4781	0.1027	0.0748	**0.5774**	**0.0895**	**0.0632**	0.5416	0.0944	0.0674
96	0.4666	0.1043	0.0751	**0.5078**	**0.0989**	**0.0708**	0.4904	0.1012	0.0726
168	**0.4453**	**0.1068**	**0.0767**	0.3738	0.1143	0.0843	0.4168	0.1100	0.0800

Table 5

Evaluation results of recurrent models with hidden state size $h = 50$ on the electricity dataset

F	TKAN			GRU			LSTM		
	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE
12	**0.2513**	0.0809	0.0631	0.1929	**0.0756**	**0.0583**	0.2130	0.1063	0.0840
24	**0.3033**	0.0940	0.0744	0.1615	0.0768	0.0593	0.1344	**0.0756**	**0.0581**
48	0.3425	0.1006	0.0791	0.0974	0.0768	0.0588	**0.3700**	**0.0763**	**0.0583**
96	**0.3733**	0.0881	0.0683	0.1886	**0.0817**	**0.0631**	0.3227	0.0915	0.0716

Table 6

Evaluation results of recurrent models with hidden state size $h = 50$ on the weather dataset

F	TKAN			GRU			LSTM		
	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE
12	0.6563	0.0773	0.0547	0.6879	0.0724	0.0497	**0.6929**	**0.0717**	**0.0484**
24	0.5880	0.0889	0.0640	0.6081	0.0858	0.0611	**0.6319**	**0.0820**	**0.0566**
48	0.4703	0.1040	0.0760	**0.5187**	**0.0981**	**0.0712**	0.4840	0.1021	0.0745
96	0.4703	0.1041	0.0748	0.4368	0.1077	0.0791	**0.4888**	**0.1016**	**0.0726**
168	0.4393	0.1075	0.0771	0.4121	0.1107	0.0807	**0.4459**	**0.1069**	**0.0767**

The experiments use an optimized KAN implementation[3], which addresses performance limitations of the original KAN framework while achieving speed comparable to MLP.

Based on Table 2, it can be concluded that KAN outperforms MLP on the selected dataset. Bold values indicate the best results, while underlined values correspond to approximately the same number of parameters. The simplest KAN configuration with a single hidden layer of 5 neurons achieves accuracy comparable to an MLP with a significantly larger number of parameters and a more complex configuration – three hidden layers with 20 neurons each. Adding 10 neurons to a single hidden layer in KAN (third row) improved accuracy more effectively than adding an additional hidden layer with 10 neurons (fourth row). Notably, the two lowest error rates were achieved with a refined spline grid. This suggests that KAN's internal degrees of freedom, implemented via splines, grant the architecture greater expressive power in approximating data dependencies.

In experiments using recurrent models (Tables 3 and 5), for energy consumption data, R^2 increases with the forecasting horizon for TKAN, while it decreases for weather data (Tables 4 and 6) and other models. At a hidden state size of $h = 50$, TKAN consistently achieves the highest R^2 across most experiments, indicating its potential for optimal energy consumption forecasting when further optimizing configurations.

To improve generalization, regularization techniques will be applied next. As shown in loss curves (e.g., $F = 96$, $h = 100$ in Fig. 2), TKAN demonstrates potential for further training, whereas LSTM and GRU exhibit overfitting, evidenced by diverging training and validation loss curves.

On weather data, TKAN demonstrates superior performance for the longest forecasting horizon when the hidden state size is set to $h = 100$. Unlike other models, the decline in the coefficient of determination (R^2) slows down starting from $F = 48$, indicating improved stability in long-term predictions. This suggests that TKAN retains greater explanatory power as the forecasting horizon increases, making

[3] GitHub – Blealtan/efficient-kan: An efficient pure-PyTorch implementation of Kolmogorov-Arnold Network (KAN), Available: https://github.com/Blealtan/efficient-kan (Accessed 08.04.2025)

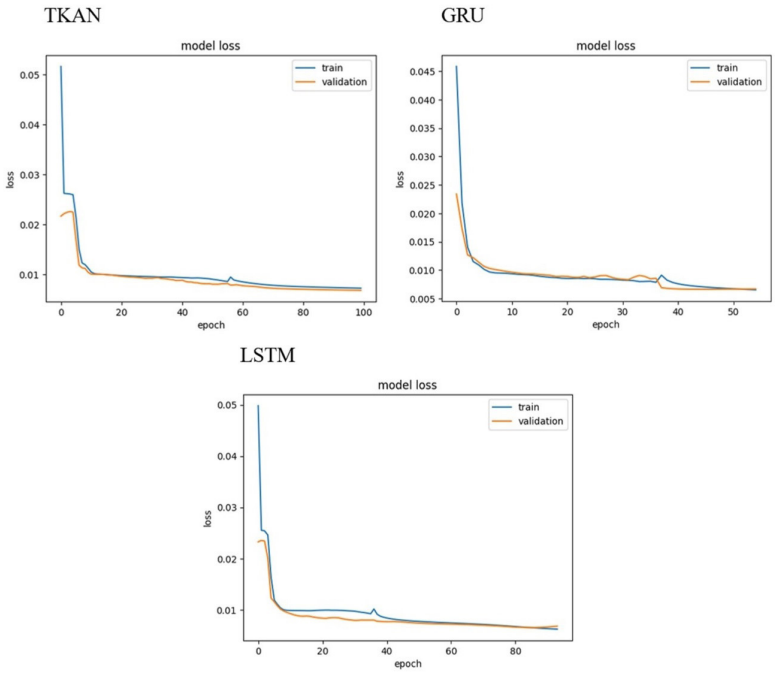


Fig. 2. Loss Functions for Recurrent Models (h = 100)

it a viable alternative to traditional architectures like LSTM and GRU for long-term time series forecasting tasks.

Enhancement through dropout integration in TKAN architecture

To mitigate overfitting and improve model robustness in time series forecasting, we propose a refined architectural framework by incorporating dropout regularization into the TKAN architecture. Dropout, a well-established technique for preventing co-adaptation of neurons during training [15], introduces stochastic deactivation of neurons, thereby enhancing generalization by reducing dependency on specific pathways.

Three variants of the modified architecture were evaluated:

1. TKAN1: A dual-dropout configuration with hidden state size h = 50, featuring dropout layers (rate = 0.1) after each recurrent layer.

2. TKAN2: A single-dropout variant with h = 40 and a dropout rate of 0.2 applied after the first recurrent layer.

3. TKAN3: A dual-dropout design with h = 40, KAN layer output size h_k = 15, and dropout rate = = 0.1 after each recurrent layer.

Experimental results validate the efficacy of this approach (Table 7), TKAN1 achieves the best performance on the longest forecasting horizon, outperforming other models. Additionally, TKAN3 demonstrates the highest R^2 for shorter-term forecasts, highlighting its adaptability to varying temporal dependencies.

This systematic integration of dropout layers underscores TKAN's capacity to harmonize structural complexity with regularization, positioning it as a competitive alternative to traditional recurrent architectures in multi-horizon forecasting tasks.

Conclusion

The empirical analysis presented on this paper demonstrates that KAN serve as a promising alternative to MLP in time series forecasting. When extended to recurrent architectures via TKAN, the

Table 7

Evaluation results of TKAN with regularization on the electricity dataset

F	TKAN			GRU			LSTM		
	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE
12	**0.3801**	0.0949	0.0729	0.3088	0.1025	0.0802	0.2775	**0.0818**	**0.0634**
24	0.3344	0.0978	**0.0755**	**0.3540**	**0.0976**	0.0762	0.3480	0.1013	0.0795
48	**0.3709**	**0.0988**	0.0770	0.3220	0.1019	0.0807	0.3693	0.0989	**0.0767**
96	**0.3483**	**0.1010**	**0.0791**	0.2877	0.1047	0.0834	0.3441	0.1012	**0.0791**

integration of LSTM-inspired gating mechanisms with spline-based parameterization exhibits superior accuracy for extended forecasting horizons while maintaining reduced architectural complexity. Specifically, proposed in the paper modification TKAN1, augmented with dual Dropout layers (dropout rate = 0.1) and a hidden state size $h = 50$, achieves the highest $R^2 = 0.3483$ and lowest RMSE 0.1010 on the Electricity dataset for extended forecasting horizons $F = 96$.

REFERENCES

1. **Holt C.C.** Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, 2004, Vol. 20, No. 1, Pp. 5–10. DOI: 10.1016/j.ijforecast.2003.09.015

2. **Winters P.R.** Forecasting sales by exponentially weighted moving averages. *Management Science*, 1960, Vol. 6, No. 3, Pp. 324–342. DOI: 10.1287/mnsc.6.3.324

3. **Box G.E.P., Jenkins G.M., Reinsel G.C., Ljung G.M.** *Time series analysis: Forecasting and control*, 5th ed. New Jersey: John Wiley & Sons, 2015.

4. **Lim B., Zohren S.** Time-series forecasting with deep learning: a survey. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 2021, Vol. 379, No. 2194, Art. no. 20200209. DOI: 10.1098/rsta.2020.0209

5. **Chigasova M.A., Maleev O.G.** Prognozirovanie kotirovok aktsii s ispol'zovaniem neironnoi seti [Stock Price Forecasting Using Neural Networks]. *Sovremennye tekhnologii v teorii i praktike programmirovaniia* [*Modern technologies in the theory and practice of programming*], 2022. S. 180–182.

6. **Bachmann G., Anagnostidis S., Hofmann T.** Scaling MLPs: A tale of inductive bias. *arXiv:2306.13575*, 2023. DOI: 10.48550/arXiv.2306.13575

7. **Liu Z., Wang Y., Vaidya S., Ruehle F., Halverson J., Soljačić M., Hou T.Y., Tegmark M.** KAN: Kolmogorov-Arnold networks. *arXiv:2404.19756*, 2024. DOI: 10.48550/arXiv.2404.19756

8. **Cybenko G.** Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 1989, Vol. 2, Pp. 303–314. DOI: 10.1007/BF02551274

9. **Fakhoury D., Fakhoury E., Speleers H.** ExSpliNet: An interpretable and expressive spline-based neural network. *Neural Networks*, 2022, Vol. 152, Pp. 332–346. DOI: 10.1016/j.neunet.2022.04.029

10. **Han X., Zhang X., Wu Y., Zhang Z., Wu Z.** Are KANs effective for multivariate time series forecasting? arXiv:2408.11306, 2024. DOI: 10.48550/arXiv.2408.11306

11. **Hewamalage H., Bergmeir C., Bandara K.** Recurrent neural networks for time series forecasting: Current status and future directions. International Journal of Forecasting, 2021, Vol. 37, No. 1, Pp. 388–427. DOI: 10.1016/j.ijforecast.2020.06.008

12. **Genet R., Inzirillo H.** TKAN: Temporal Kolmogorov-Arnold networks. *arXiv:2405.07344*, 2024. DOI: 10.48550/arXiv.2405.07344

13. **Hochreiter S., Schmidhuber J.** Long short-term memory. *Neural Computation*, 1997, Vol. 9, No. 8, Pp. 1735–1780. DOI: 10.1162/neco.1997.9.8.1735

14. **Cho K., van Merriënboer B., Gulcehre C., Bahdanau D., Bougares F., Schwenk H., Bengio Y.** Learning phrase representations using RNN encoder–decoder for statistical machine translation. *arXiv:1406.1078*, 2014. DOI: 10.48550/arXiv.1406.1078

15. **Gal Y., Ghahramani Z.** A theoretically grounded application of dropout in recurrent neural networks. *arXiv:1512.05287*, 2015. DOI: 10.48550/arXiv.1512.05287

INFORMATION ABOUT AUTHORS / СВЕДЕНИЯ ОБ АВТОРАХ

Oleg G. Maleev
Малеев Олег Геннадьевич
E-mail: olegmg@bk.ru
ORCID: https://orcid.org/0000-0001-7810-7973

Olga A. Kovaleva
Ковалева Ольга Анатольевна
E-mail: kovaleva.oa@edu.spbstu.ru

Submitted: 01.07.2025; Approved: 25.11.2025; Accepted: 26.11.2025.
Поступила: 01.07.2025; Одобрена: 25.11.2025; Принята: 26.11.2025.

Computing, Telecommunication and Control, 2025, Vol. 18, No. 4, Pp. 30–43.
Информатика, телекоммуникации и управление. 2025. Том 18, № 4. С. 30–43.

Research article
DOI: https://doi.org/10.18721/JCSTCS.18403
UDC 004.8

GENERATIVE ADVERSARIAL NETWORK FOR CLASSIFICATION OF MECHANICAL FAULT DIAGNOSIS MODEL

C. Guo ✉ , *V.V. Potekhin*

Peter the Great St. Petersburg Polytechnic University,
St. Petersburg, Russian Federation

✉ chenxiguo.academic@gmail.com

Abstract. The scarcity and imbalance of annotated fault data pose significant challenges to the reliability of intelligent industrial diagnostics. To address this issue, we propose an integrated fault diagnosis framework based on multi-domain feature fusion and generative adversarial networks (GANs). Unlike traditional approaches that treat generation and classification as independent stages, our model unifies these two processes. This method achieves diagnosis by transforming raw vibration signals into multi-domain representations (time domain, frequency domain, and time-frequency domain). The core innovation lies in the restructured generator architecture: a Transformer encoder captures global signal correlations, combined with an Efficient Channel Attention (ECA) mechanism for adaptive recalibration of feature weights, ensuring high-fidelity sample synthesis. Additionally, the model employs a dual-function discriminator that distinguishes genuine from synthetic samples while directly performing multi-class fault classification. Extensive experiments on CWRU and JNU benchmark datasets demonstrate that this approach surpasses existing state-of-the-art algorithms, achieving superior performance in Structural Similarity (SSIM), Peak Signal-to-Noise Ratio (PSNR), and diagnostic accuracy. This end-to-end solution effectively mitigates data scarcity challenges in industrial settings.

Keywords: fault diagnosis, generative adversarial networks, limited data, supervised learning, time-series analysis

Acknowledgements: The research was financially supported by the China Scholarship Council (CSC.202309810002).

Citation: Guo C., Potekhin V.V. Generative adversarial network for classification of mechanical fault diagnosis model. Computing, Telecommunications and Control, 2025, Vol. 18, No. 4, Pp. 30–43. DOI: 10.18721/JCSTCS.18403

Научная статья
DOI: https://doi.org/10.18721/JCSTCS.18403
УДК 004.8

ГЕНЕРАТИВНО-СОСТЯЗАТЕЛЬНАЯ СЕТЬ ДЛЯ КЛАССИФИКАЦИИ МОДЕЛИ ДИАГНОСТИКИ МЕХАНИЧЕСКИХ НЕИСПРАВНОСТЕЙ

Ч. Го ✉ , *В.В. Потехин*

Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Российская Федерация

✉ chenxiguo.academic@gmail.com

Аннотация. Нехватка и несбалансированность аннотированных данных о неисправностях создают серьезные проблемы для надежности интеллектуальной промышленной диагностики. Для решения этой проблемы мы предлагаем интегрированную систему диагностики неисправностей, основанную на слиянии многодоменных характеристик и генеративных состязательных сетях (GAN). В отличие от традиционных подходов, которые рассматривают генерацию и классификацию как независимые этапы, наша модель объединяет эти два процесса. Этот метод позволяет проводить диагностику путем преобразования необработанных сигналов вибрации в многодоменные представления (временная область, частотная область и временная-частотная область). Основная инновация заключается в реструктурированной архитектуре генератора: кодер Transformer улавливает глобальные корреляции сигналов в сочетании с механизмом Efficient Channel Attention (ECA) для адаптивной перекалибровки весов признаков, обеспечивая высокую точность синтеза образцов. Кроме того, модель использует дискриминатор с двойной функцией, который отличает подлинные образцы от синтетических, одновременно выполняя многоклассовую классификацию неисправностей. Обширные эксперименты на эталонных наборах данных CWRU и JNU демонстрируют, что этот подход превосходит существующие современные алгоритмы, достигая превосходных результатов по структурному сходству (SSIM), пиковому отношению сигнал/шум (PSNR) и точности диагностики. Это комплексное решение эффективно смягчает проблемы нехватки данных в промышленных условиях.

Ключевые слова: диагностика неисправностей, генеративно-состязательные сети, ограниченные данные, контролируемое обучение, анализ временных рядов

Финансирование: Исследование выполнено при финансовой поддержке Китайского государственного комитета по стипендиям (CSC.202309810002).

Для цитирования: Guo C., Potekhin V.V. Generative adversarial network for classification of mechanical fault diagnosis model // Computing, Telecommunications and Control. 2025. T. 18, № 4. С. 30–43. DOI: 10.18721/JCSTCS.18403

Introduction

As was stated in our previous article [22], with the rapid advancement of Industry 4.0, the demand for the health monitoring and operational stability of intelligent industrial equipment has significantly increased. As core components of rotating machinery, the condition of rolling bearings directly impacts the safety and stability of industrial systems. Due to their operation under variable speed and load conditions, bearings are susceptible to a variety of faults [22].

Traditional fault diagnosis methods typically rely on the analysis of 1D vibration signals from sensors, requiring professionals to manually extract and assess signal features before designing classifiers. Although these approaches have achieved high accuracy, they are often time-consuming and

labor-intensive, rendering them ill-suited for the automated, intelligent diagnostic requirements of modern manufacturing systems [22].

In recent years, data-driven methodologies, particularly convolutional neural networks (CNNs), have emerged as powerful tools for capturing nonlinear fault characteristics without human intervention. However, the efficacy of these deep learning models is often contingent upon the availability of massive, annotated datasets. In practical industrial settings, acquiring high-quality fault data is challenging, leading to issues of data paucity and class imbalance. Consequently, standard models frequently fail to generalize or maintain high diagnostic accuracy when training samples are scarce [6, 7].

To mitigate the challenges of data scarcity, generative adversarial networks (GANs) [8] have been adopted as a robust strategy for data augmentation. By synthesizing realistic fault samples, GANs can rebalance datasets and enhance model robustness [9]. It is important to note that while our previous research explored the use of improved vision transformers (ViT) as standalone classifiers for fault diagnosis [22], the current study focuses on a different architectural approach. Specifically, rather than relying on an external ViT classifier, this work aims to optimize the generative process itself to produce higher fidelity samples for an integrated diagnostic framework.

The primary objective of this research is to develop a deep learning framework tailored for bearing fault diagnosis under severely limited data conditions. We propose an integrated model that fuses multi-domain features (time, frequency and time-frequency). Distinct from prior approaches, the novelty of this model lies in its generator architecture, which incorporates a Transformer encoder to capture global signal interactions and an Efficient Channel Attention (ECA) mechanism to refine feature representation.

The core challenge addressed in this study is the generation of high-fidelity synthetic data to compensate for the lack of training samples. By strategically mixing original and generated data, we aim to construct an extended, balanced dataset that serves as the foundation for highly accurate fault classification performed directly by the model's integrated discriminator.

The specific contributions of this paper are summarized as follows:

1. **Investigation of data scarcity:** We conduct a systematic evaluation of bearing fault diagnosis performance under conditions of limited and unbalanced training samples using public datasets.

2. **Novel data enhancement model:** We develop a multi-domain feature fusion GAN. This model uniquely integrates features from three domains and utilizes adversarial learning with attention mechanisms to ensure the generation of synthetic data that closely mimics the physical properties of real fault signals.

3. **Performance benchmarking:** We perform a comprehensive comparative analysis against state-of-the-art algorithms. The experimental results validate that the proposed method achieves superior reliability and applicability for industrial fault diagnosis tasks.

Related works

Traditional and deep learning-based fault diagnosis methods

Bearing fault diagnosis techniques have evolved from traditional statistical methods to advanced methods based on deep learning. Traditional methods, such as Gaussian Mixture Models (GMMs) and Hidden Markov Models (HMMs), rely heavily on hand-extracted features, such as Mel Frequency Cepstral Coefficients (MFCCs), for modeling speech features. However, these methods have limited performance in dealing with high-dimensional data and complex working conditions.

With the rise of deep learning, models such as CNN, recurrent neural networks (RNNs) and long-short-term memory networks (LSTMs) are widely used in fault diagnosis. These models can automatically learn the hierarchical features of the original vibration signals, improving the accuracy and robustness of fault identification.

In [20], a transformer-based conditional generative adversarial network migration learning model was proposed, which enhances the quality and diversity of the generated data by introducing sample labeling information, thus improving the performance of cross-domain fault diagnosis.

In [21], a bearing fault diagnosis study based on a multimodal approach combined with a multi-scale time-frequency and statistical feature fusion model was proposed, which is able to better handle non-stationary and nonlinear vibration data.

Limitations

Although the above methods have made significant progress in bearing fault diagnosis, there are still some challenges. Traditional methods have limited performance in dealing with high-dimensional data and complex working conditions, while deep learning methods face problems such as high consumption of computational resources and strong dependence on a large amount of labeled data.

Rationale for GAN

GANs provide a compelling solution to key challenges in bearing fault diagnosis, especially data scarcity and class imbalance. By introducing an adversarial framework between the generator and the discriminator, GANs can learn complex data distributions and synthesize real fault samples to expand limited data sets. This is particularly important in industrial environments where access to labeled fault data is costly or impractical. The generator captures subtle fault features-often from the time, frequency, or time-frequency domain-while the discriminator ensures sample quality through adversarial training. This dynamic change not only improves the robustness of diagnostic models, but also enhances their generalizability to real-world scenarios. Thus, GANs are a promising direction for building more accurate and resilient fault diagnosis systems under constrained data conditions.

Materials and methods

This study analyzes various bearing operating states using the Case Western Reserve University (CWRU) dataset [10]. This dataset provides information on the vibration characteristics of bearings in different states including normal operation, inner ring, outer ring and ball damage.

Fig. 1 shows a schematic of a bearing including the main structural elements: inner ring, outer ring and balls. Damage can be associated with different surface areas or different defect sizes, resulting in differences in bearing operating conditions. These differences create a complex classification problem that requires the use of deep learning techniques to accurately diagnose faults.

Data preprocessing involves converting vibration signals into multimodal representations, which allows the consideration of time, frequency, and time-frequency characteristics for further analysis and model training.

In the CWRU dataset, bearing damage is categorized by type and location of occurrence. The main categories include: ball damage, inner ring damage, outer ring damage, and normal condition. Additionally, damage is differentiated by diameter, which is represented by values of 0.007, 0.014, 0.021 etc. Thus, by combining the different types of damage and their diameters, the operational condition of bearings can be divided into ten categories.

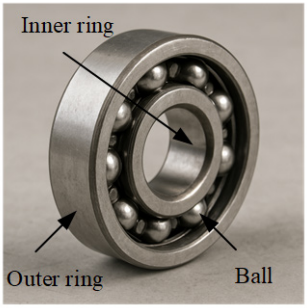


Fig. 1. Bearing diagram

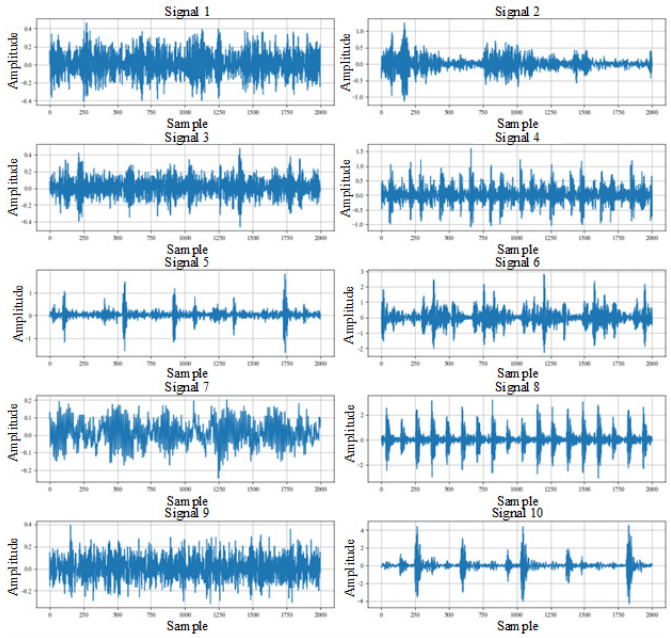


Fig. 2. Vibration signals of bearings in different damage states

In general, bearing vibration signals recorded under different operating conditions have markedly different characteristics, as evidenced by the data presented in Fig. 2. For example, in the range from sample 0 to sample 1000, signal 2 exhibits almost no pronounced vibration, while the vibration of signal 8 has a distinctly periodic character.

Furthermore, in the range from sample 200 to sample 1200, there is a significant difference in the vibration amplitudes of signal 1 and signal 9. These differences illustrate the complexity of analyzing bearing condition data and confirm the need for methods capable of efficiently processing and analyzing data with such variations.

As shown in Fig. 3, in order to obtain a complete set of vibration signal characteristics, this study used a time domain data transformation approach to transform the data into different representations.

The sample data were transformed in three different domains:

1. Time domain – the original signals in their original form.

2. Frequency domain – transforming the data using Fast Fourier transform (FFT) to reveal the frequency components of the signal.

3. Time-frequency domain – a representation obtained using a time-frequency domain transform, such as the Short-time Fourier transform (STFT), which allows you to analyze the dynamics of frequency components over time.

This approach provides a comprehensive analysis of signal characteristics, which is a key step for successful model training and bearing fault diagnosis.

Design of bearing fault diagnostic modeling

To address the persistent challenges of bearing fault diagnosis under data-constrained and imbalanced conditions, this study introduces a unified, end-to-end deep learning framework. Unlike our previous work [22], which relied on an external classifier coupled with a generative model, the approach proposed in this study orchestrates sample generation and fault diagnosis within a single, cohesive architecture. As illustrated in Fig. 4, the framework is designed to synthesize high-fidelity multi-domain samples while simultaneously executing precise 10-category fault diagnosis through an integrated mechanism.

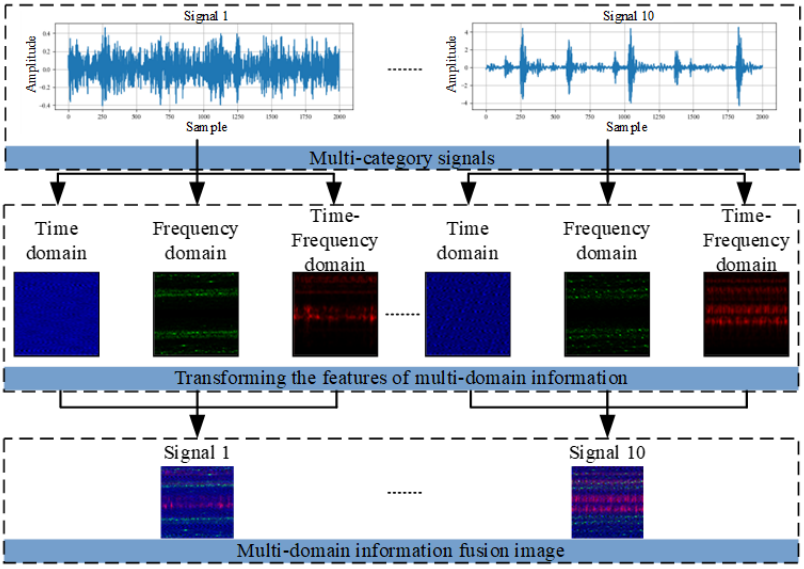


Fig. 3. Information about vibration signals in time, frequency and frequency-time domains

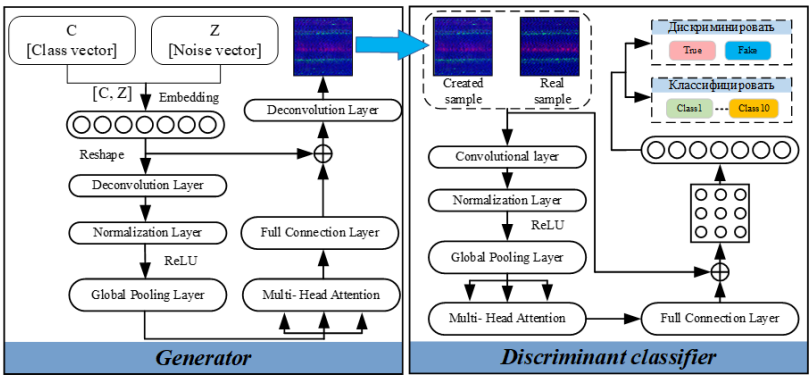


Fig. 4. Proposed fault diagnosis model

Optimization of hyperparameters is pivotal for ensuring model convergence and diagnostic reliability. Through empirical validation, the framework employs an Adam optimizer initialized with a learning rate of 0.001. Training is conducted with a batch size of 128 over 6000 epochs to guarantee stable feature extraction and distribution matching.

The primary function of the generator is to map latent noise distributions into interpretable, multi-domain signal representations. The generative process is initiated by fusing two distinct input vectors:

– Latent Vector (Z): A noise vector sampled from a standard Gaussian distribution to induce sample diversity.

– Label Embedding (C): A category-specific vector projected into a high-dimensional space via an embedding layer to condition the generation.

Mathematically, the embedding transformation is governed by the weight matrix:

$$\mathbf{W}_{\text{embed}} \in \mathbb{R}^{n_{\text{classes}} \times d_{\text{embed}}}, \quad (1)$$

where n_{class} represents the number of fault categories and d_{embed} denotes the embedding dimension. The vectors Z and C are fused via element-wise multiplication and subsequently reshaped into a four-dimensional tensor, serving as the foundational input for the deep deconvolutional layers.

To capture long-range dependencies within the feature maps $X \in \mathbb{R}^{B\times C\times H\times W}$, a Transformer-based enhancement module is embedded within the generator. Initially, spatial information is compressed into a channel descriptor Y via Global Average Pooling (GAP):

$$Y_c = \frac{1}{H\times W}\sum_{i=1}^{H}\sum_{j=1}^{W} X_{c,i,j}, \tag{2}$$

where c denotes the channel index, while H and W represent spatial dimensions. The resulting channel descriptor Y_c is flattened into a sequence to facilitate processing by the Multi-Head Attention (MHA) mechanism [7]. The MHA module dynamically models inter-channel correlations by computing Query (Q), Key (K), and Value (V) matrices through learned linear projections:

$$\text{head}_i = f_{\text{softmax}}\left(\frac{Q_i K_i^T}{\sqrt{d_k}}\right)V_i; \tag{3}$$

$$\text{MHA}\left(Q,K,V\right) = \text{Concat}\left(\text{head}_1,\ldots,\text{head}_h\right)W, \tag{4}$$

where d_k is the dimension of the attention heads, h is the number of heads, and $W \in R^{D\times D}$ is the linear projection matrix. This mechanism allows the generator to contextualize local features within the global signal structure. Prior to attention computation, a Normalization Layer (NL) is applied to stabilize gradients:

$$Y' = NL\left(Y + \text{MultiHead}\left(Q,K,V\right)\right), \tag{5}$$

where Y and Y' represent the embedding sequence and MHA output, respectively, and $NL(\cdot)$ denotes the normalization operation.

To further refine feature saliency, an ECA mechanism is integrated. The ECA module adaptively recalibrates channel weights, emphasizing informative features while suppressing noise. The re-weighted feature output $\mathbf{O}_{\text{ECA}}$ is computed as:

$$\mathbf{O}_{\text{ECA}} = \sigma\left(\mathbf{Y}'\right)\cdot\mathbf{X}, \tag{6}$$

where $\sigma(\cdot)$ denotes the sigmoidal activation function and $\mathbf{O}_{\text{ECA}}$ is the output feature after weighting the channel attention.

A defining characteristic of this framework is the dual-function discriminator. Distinct from methods that utilize separate downstream classifiers (e.g., [22]), the discriminator in this model is engineered to directly perform multi-class fault diagnosis alongside its adversarial duties.

The discriminator loss function consists of two components: classification loss and category classification loss. For real samples S_{real}, the loss for real samples $S_{D,\text{real}}$ is defined as:

$$L_{D,\text{real}} = -\mathbb{E}_{S_{\text{real}}\sim P_{\text{data}}}\left[\log D\left(S_{\text{real}}\right)\right] + \mathbb{E}_{y_{\text{real}}}\left[\log P\left(y = y_{\text{real}}\left|S_{\text{real}}\right.\right)\right], \tag{7}$$

where y_{real} represents the true label.

For S_g generated samples, the generated loss $L_{D,\text{fake}}$ is defined as:

$$L_{D,\text{fake}} = \mathbb{E}_{z\sim P_z}\left[\log D\left(S_g\right)\right] - \mathbb{E}_{y_{\text{gen}}}\left[\log P\left(y = y_{\text{gen}} \middle| S_g\right)\right], \tag{8}$$

where y_{gen} denotes the generated label.

To improve the robustness of training, a Gradient Penalty (GP) term is introduced:

$$L_{\text{GP}} = \lambda_{\text{GP}} \cdot \mathbb{E}_{\hat{S}\sim P_S}\left[\left(\left\|\nabla_{\hat{S}} D\left(\hat{S}\right)\right\|_2 - 1\right)^2\right], \tag{9}$$

where P_S is the uniform sampling distribution between real and generated samples, and λ_{GP} is the regularization parameter. Therefore, the total discriminator loss is expressed as:

$$L_D = L_{D,\text{real}} + L_{D,\text{fake}} + L_{\text{GP}}. \tag{10}$$

The task of the generator is to fool the discriminator while ensuring that the generated samples are assigned the correct class labels. The generator loss function is defined as:

$$L_{\text{G}} = -\mathbb{E}_{z\sim P_z}\left[\log D\left(S_g\right)\right] + \mathbb{E}_{y_{\text{gen}}}\left[\log P\left(y = y_{\text{gen}} \middle| S_g\right)\right]. \tag{11}$$

The fault classification model achieves accurate recognition of different fault modes by supervised learning and is optimized using a loss function based on cross entropy. To further improve the performance of the model, the real and generated data are combined for training. The loss function of the fault classification model is defined as:

$$L_{total} = \min_G \max_D \left(L_D + L_G\right). \tag{12}$$

Results and discussion

In order to evaluate the model's ability to generate data, this paper uses a joint sample quality assessment method to evaluate the performance of the generated samples. The method includes PSNR and SSIM [12, 13, 22].

PSNR measures the total pixel error between the generated image and the original image. PSNR is calculated as follows:

$$\text{MSE} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(x_{ij} - y_{ij}\right)^2; \tag{13}$$

$$\text{PSNR} = 20 \cdot \log_{10}\left(\frac{I_{\text{MAX}}}{\sqrt{\text{MSE}}}\right), \tag{14}$$

where MSE is the mean square error; M and N are the width and height of the image, respectively; X_{ij} and Y_{ij} are the pixel values of the original and generated images, respectively; I_{MAX} is the maximum possible pixel value.

SSIM evaluates the perceptual similarity between the generated image and the original image. Its formula is as follows:

$$\mathrm{SSIM}(x,y)=\frac{(2\mu_x\mu_y+C_1)(2\sigma_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)}, \tag{15}$$

where μ_x and μ_y are the mean values of the two images; μ_x^2 and μ_y^2 are their variance; σ_{xy} is the covariance of the two images, respectively; C_1 and C_2 are constants to ensure computational stability.

Fig. 5 shows the variation of the training loss of the proposed model and the validation accuracy of the discriminator. From it, it is evident that the adversarial training of the generator and discriminator gradually stabilizes after 1000 epochs. The discriminator accuracy is also close to 100%.

The Figs. 6 and 7 show the quality scores of different models for each category of generated samples, including GAN [8], ACGAN [14], DCGAN [15] and the method used in this paper. Comparing the histograms for each category shows that the method proposed in this paper outperforms its counterparts in terms of average SSIM and PSNR and ranks first in all ten categories.

In addition, the difference between the SSIM and PSNR scores for each category is only ±0.02, indicating the high stability and accuracy of the model in training the features in all categories. These results confirm that the proposed method can effectively meet the challenge of generating high-quality data and contribute to the improvement of bearing fault diagnosis.

To further validate the effectiveness of the proposed classification method, we compared the developed model with various state-of-the-art classification models including Random Forest (RF) [16], Support Vector Machine (SVM) [17], Hierarchical CNN (H-CNN) [18], 2D-CNN [19] and the method proposed in this paper. The results of the comparative analysis are presented in Fig. 8.

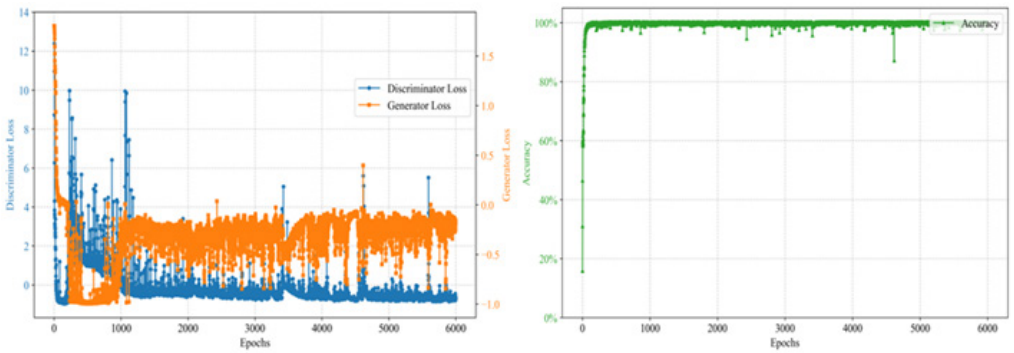


Fig. 5. Variation of losses during model training and discriminator validation accuracy

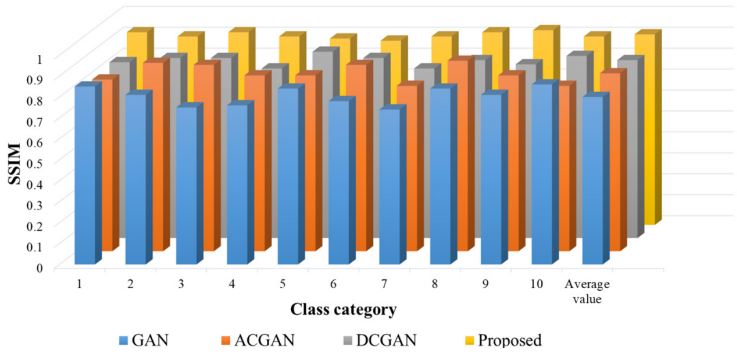


Fig. 6. Comparison of SSIM with different generation models

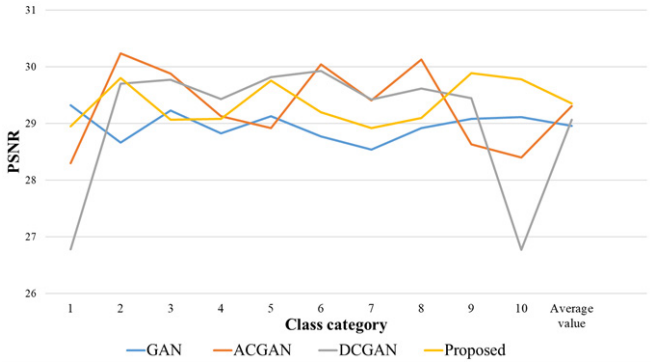


Fig. 7. Comparison of PSNR with different generation models

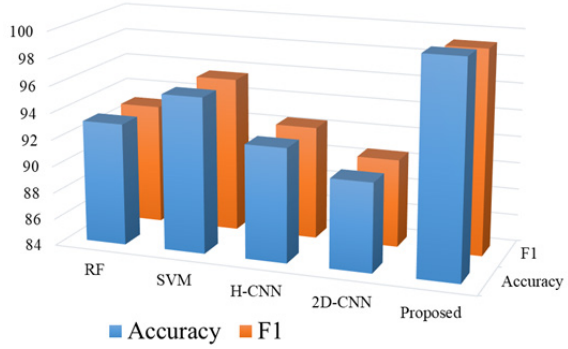


Fig. 8. Comparison of accuracy and F1 between different models

These results demonstrate the superiority of the proposed model in bearing fault diagnosis tasks, especially under conditions of limited and unbalanced data. On key classification metrics, the proposed method consistently outperforms alternative approaches, confirming its robustness and practical applicability.

Among machine learning models, the random forest and support vector machine algorithms show similar results, with classification accuracies of 93.13% and 95.63%, and F1-scores of 93.01% and 95.60%, respectively. Although SVM outperforms RF processing of high-dimensional data, its performance in the task of fault signal classification remains limited. This is due to the inherent weaknesses of traditional methods in extracting features and adapting them to high-dimensional data, which does not fully reveal the underlying features of the signals.

Among deep learning models, the method proposed in this paper demonstrates the highest performance on all metrics. The classification accuracy reaches 99.91% and the F1-score reaches 99.25%. These metrics emphasize the significant advantages of the developed model for fault classification tasks. The generation of high-quality augmented data using the proposed approach significantly improves the generalization ability and classification performance of the model, unlocking its full potential in complex diagnosis tasks.

To deeply evaluate the classification performance of the proposed model on the CWRU bearing dataset, this paper conducts relevant experiments and constructs a confusion matrix based on the test set samples to clearly demonstrate the model's ability to recognize different types of faults. Fig. 9 shows the confusion matrix, where the horizontal axis shows the fault types predicted by the model and the vertical axis shows the actual fault types. As can be seen from Fig. 9, the proposed model achieves

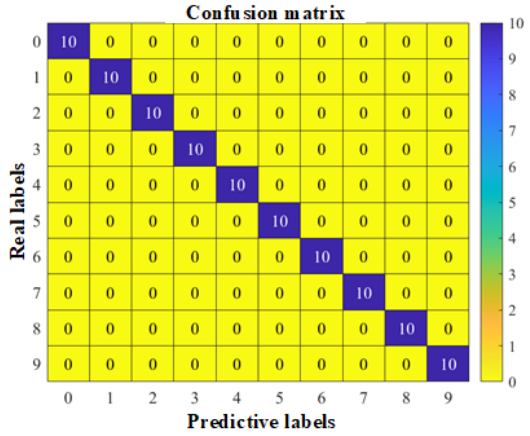


Fig. 9. Confusion matrix results

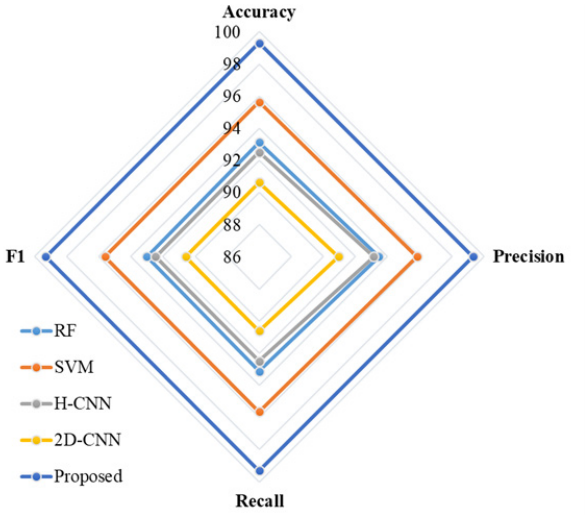


Fig. 10. The radar plot of data comparison with models

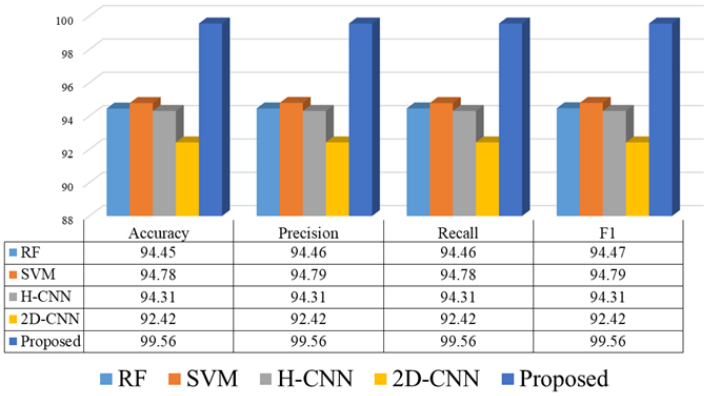
high classification accuracy of ten fault types in the CWRU dataset, which fully confirms its effectiveness and robustness in fault diagnosis tasks.

In addition, this paper further verifies the bearing data from Jinan University. In this paper, we classify the signal data in different operating conditions, and there are eight kinds of bearing data. The comparison of experimental data with different comparison models is shown in Fig. 10.

As can be seen from the data distribution graph in Fig. 10, the method proposed in this paper outperforms the compared methods in terms of accuracy, predictive value, recall and F1-score, which proves the strong generalization ability of the proposed method.

In addition, in order to verify the generalization ability of the proposed model in this paper, experiments are conducted on the JNU dataset, and the experimental results of the comparison models are shown in Fig. 11. As can be seen from the figure, the proposed model exhibits the best diagnostic results.

In summary, the experiments on CWRU and JNU datasets show that the model proposed in this paper is effective in diagnosing the operating conditions under different mechanical conditions.



	Accuracy	Precision	Recall	F1
RF	94.45	94.46	94.46	94.47
SVM	94.78	94.79	94.78	94.79
H-CNN	94.31	94.31	94.31	94.31
2D-CNN	92.42	92.42	92.42	92.42
Proposed	99.56	99.56	99.56	99.56

Fig. 11. Comparing the results of the model on the JNU dataset

Conclusion

This study presents a novel deep learning-based approach for bearing fault diagnosis, tailored to address the data scarcity and diagnostic complexity characteristic of the industry 4.0 context. By integrating an enhanced GAN with a multi-domain feature fusion framework, the proposed method effectively augments limited datasets and improves diagnostic performance. The key findings are summarized as follows:

- Data augmentation: The improved GAN demonstrates superior capability in generating high-fidelity samples, as evidenced by higher PSNR and SSIM values compared to conventional data generation methods.
- Multi-domain feature learning: By leveraging temporal, frequency, and time-frequency domain features, and enhancing them through attention mechanisms, the model achieves robust and comprehensive feature representation.
- Superior classification accuracy: The proposed model achieves a diagnostic accuracy of 99.91%, significantly outperforming established baselines such as SVM and 2D-CNN.

Overall, this work contributes to the advancement of intelligent diagnostic systems by providing a scalable and generalizable solution. Future research may explore extending the framework to multi-task learning scenarios and applying it to other complex industrial systems for enhanced prognostic capabilities.

REFERENCES

1. **Buinosov A.P., Vasiliev V.A., Baitov A.S., Ivanov A.I.** Bench testing of algorithms for diagnosing rolling bearings of the on-board diagnostic system and forecasting the service life of the main and auxiliary components of the MCRSU. *Herald of the Ural State University of Railway Transport*, 2021, Vol. 51, No. 3, Pp. 40–49. DOI: 10.20291/2079-0392-2021-3-40-49

2. **Gerike P.B., Yeshcherkin P.V.** Development of a unified diagnostic criterion for mining shovel power-mechanical equipment rolling bearings' diagnosis. *Vestnik nauchnogo tsentra po bezopasnosti rabot v ugol'noi promyshlennosti* [*Bulletin of the Scientific Center for Safety in the Coal Industry*], 2020. No. 2, Pp. 65–72.

3. **Xu X., Yang X., He C., Shi P., Hua C.** Adversarial domain adaptation model based on LDTW for extreme partial transfer fault diagnosis of rotating machines. *IEEE Transactions on Instrumentation and Measurement*, 2024, Vol. 73, Art. no. 3538811. DOI: 10.1109/TIM.2024.3476708

4. **Huang F., Zhang K., Zheng Q., Li Z., Lai X., Ding G., Zhao M.** An open-set method for diagnosing unknown bogie bearing faults using a hybrid open score relation network. *IEEE Transactions on Instrumentation and Measurement*, 2024, Vol. 73, Art. no. 3539211. DOI: 10.1109/TIM.2024.3480217

5. **Triandofilidi I., Kotsyuba I.Yu.** Prediction the operating state of an electric vehicle charging station using neural networks. *Modern High Technologies*, 2024, No. 4, Pp. 101–106. DOI: 10.17513/snt.39978

6. **Romanovskaya Y.M., Ilyushin E.A.** Sound augmentation methods. *International Journal of Open Information Technologies*, 2023, Vol. 11, No. 3, Pp. 12–18.

7. **Ibryaeva O.L., Mohammad M.N.** Fault diagnosis of rolling bearings using spectral peaks and neural networks. *Bulletin of the South Ural State University. Series: Computational Mathematics and Software Engineering*, 2022, Vol. 11, No. 2, Pp. 59–71. DOI: 10.14529/cmse220205

8. **Goodfellow I., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y.** Generative adversarial networks. *Communications of the ACM*, 2020, Vol. 63, No. 11, Pp. 139–144. DOI: 10.1145/3422622

9. **Uzkikh G.I.** Application of generative adversarial networks (GANs) in image processing. *Vestnik nauki* [*Science Herald*], 2024, Vol. 77, No. 8 (4), Pp. 182–185.

10. **Zou Y., Zhao W., Liu Y., Zhang Z.** Research on rolling bearing fault diagnosis method based on the combination of optimized VMD and BP neural network. *Bulletin of PNU*, 2023, Vol. 70, No. 3, Pp. 41–56.

11. **Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A.N., Kaiser L., Polosukhin I.** Attention is all you need. *arXiv:1706.03762*, 2017. DOI: 10.48550/arXiv.1706.03762

12. **Lian J., Yang Z., Liu J., Sun W., Zheng L., Du X., Yi Z., Shi B., Ma Y.** An overview of image segmentation based on pulse-coupled neural network. *Archives of Computational Methods in Engineering*, 2021, Vol. 28, Pp. 387–403. DOI: 10.1007/s11831-019-09381-5

13. **Guo C., Lian J., Kang Y., Zheng L., Zhang M.** The limited-FCMSPCNN algorithm for image defogging. *Journal of Physics: Conference Series*, 2021, Vol. 1961, Ar. no. 012024. DOI: 10.1088/1742-6596/1961/1/012024

14. **Odena A., Olah C., Shlens J.** Conditional image synthesis with auxiliary classifier GANs. *Proceedings of the 34th International Conference on Machine Learning*, 2017, Vol. 70, Pp. 2642–2651. DOI: 10.5555/3305890.3305954

15. **Radford A., Metz L., Chintala S.** Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv:1511.06434*, 2015. DOI: 10.48550/arXiv.1511.06434

16. **Paul A., Mukherjee D.P., Das P., Gangopadhyay A., Chintha A.R., Kundu S.** Improved random forest for classification. *IEEE Transactions on Image Processing*, 2018, Vol. 27, No. 8, Pp. 4012–4024. DOI: 10.1109/TIP.2018.2834830

17. **Cervantes J., Garcia-Lamont F., Rodríguez-Mazahua L., Lopez A.** A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*, 2020, Vol. 408, Pp. 189–215. DOI: 10.1016/j.neucom.2019.10.118

18. **Zhu X., Bain M.** B-CNN: Branch convolutional neural network for hierarchical classification. *arXiv:1709.09890*, 2017. DOI: 10.48550/arXiv.1709.09890

19. **Bird J.J., Faria D.R., Manso L.J., Ayrosa P.P.S., Ekárt A.** A study on CNN image classification of EEG signals represented in 2D and 3D. *Journal of Neural Engineering*, 2021, Vol. 18, No. 2, Art. no. 026005. DOI: 10.1088/1741-2552/abda0c

20. **Li X., Ma J., Wu J., Li Z., Tan Z.** Transformer-based conditional generative transfer learning network for cross domain fault diagnosis under limited data. *Scientific Reports*, 2025, Vol. 15, Art. no. 6836. DOI: 10.1038/s41598-025-91424-y

21. **Chen H., Li S., Lu X., Zhang Q., Zhu J., Lu J.** Research on bearing fault diagnosis based on a multimodal method. *Mathematical Biosciences and Engineering*, 2024, Vol. 21, No. 12, Pp. 7688–7706. DOI: 10.3934/mbe.2024338

22. **Guo C., Potekhin V.V., Li P., Kovalchuk E.A., Lian J.** MDFT-GAN: A multi-domain feature transformer GAN for bearing fault diagnosis under limited and imbalanced data conditions. *Applied Sciences*, 2025, Vol. 15, Art. no. 6225. DOI: 10.3390/app15116225

INFORMATION ABOUT AUTHORS / СВЕДЕНИЯ ОБ АВТОРАХ

Chenxi Guo
Го Чэньси
E-mail: chenxiguo.academic@gmail.com
ORCID: https://orcid.org/0000-0002-1497-2893

Vyacheslav V. Potekhin
Потехин Вячеслав Витальевич
E-mail: Slava.Potekhin@spbstu.ru
ORCID: https://orcid.org/0000-0001-9850-9558

Submitted: 08.05.2025; Approved: 08.12.2025; Accepted: 08.12.2025.
Поступила: 08.05.2025; Одобрена: 08.12.2025; Принята: 08.12.2025.

Computing, Telecommunication and Control, 2025, Vol. 18, No. 4, Pp. 44–52.
Информатика, телекоммуникации и управление. 2025. Том 18, № 4. С. 44–52.

Research article
DOI: https://doi.org/10.18721/JCSTCS.18404
UDC 004.03

EDUCATION SYSTEM OF ENGINEERS AND SCIENTIFIC PERSONNEL IN THE SPHERE OF QUANTUM INFORMATION TECHNOLOGIES, QUANTUM ROBUST ARTIFICIAL INTELLIGENCE AND QUANTUM RESILIENCE

V.Y. Skiba ✉, *S.A. Petrenko*, *E.M. Abakumov*

Sirius University of Science and Technology, "Sirius" Federal Territory, Krasnodar Krai, Russian Federation

✉ vskiba69@mail.ru

Abstract. Currently, quantum technologies are the driving force behind technological progress (Industry 4.0/5.0/6.0 or Society 6.0). The development of quantum information technologies gives rise to new quantum threats to information security. The requirements of the federal project "Personnel for the Digital Economy or the Data Economy" for the key information technology – "Quantum Technologies" (Priority 1) – can be fulfilled only by developing an appropriate training system for engineering and scientific personnel in the fields of quantum information technologies, quantum robust artificial intelligence and quantum resilience. The creation of a vertically integrated education system, which includes not only basic higher education, but also the training of scientific personnel and upskilling for already employed professionals, will contribute to the development of innovative technologies in general. The paper presents the results of joint efforts of the Sirius University of Science and Technology together with the Quantum Consortium of Business and Universities to train personnel in the sphere of quantum information technologies for the development of such an education system.

Keywords: quantum computing, quantum information technology, information security, educational program, new quantum security threat, quantum and post-quantum cryptography, quantum robust artificial intelligence

Acknowledgements: The results were obtained with the financial support of the project "Technologies for countering previously unknown quantum cyber threats", implemented within the framework of the state program of the "Sirius" Federal Territory "Scientific and technological development of the "Sirius" Federal Territory" (Agreement No. 23-03 dated September 27, 2024).

Citation: Skiba V.Y., Petrenko S.A., Abakumov E.M. Education system of engineers and scientific personnel in the sphere of quantum information technologies, quantum robust artificial intelligence and quantum resilience. Computing, Telecommunications and Control, 2025, Vol. 18, No. 4, Pp. 44–52. DOI: 10.18721/JCSTCS.18404

Научная статья
DOI: https://doi.org/10.18721/JCSTCS.18404
УДК 004.03

СИСТЕМА ОБРАЗОВАНИЯ ИНЖЕНЕРОВ И НАУЧНЫХ КАДРОВ В СФЕРЕ КВАНТОВЫХ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, КВАНТОВО-СИЛЬНОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И КВАНТОВОЙ УСТОЙЧИВОСТИ

В.Ю. Скиба ✉, *С.А. Петренко*, *Е.М. Абакумов*

Научно-технологический университет «Сириус», федеральная территория «Сириус», Краснодарский край, Российская Федерация

✉ vskiba69@mail.ru

Аннотация. В настоящее время квантовые технологии являются двигателем технического прогресса (промышленности 4.0/5.0/6.0 или общества 6.0). Развитие квантовых информационных технологий порождает новую квантовую угрозу информационной безопасности. Требования федерального проекта «Кадры для цифровой экономики или экономики данных» к ключевой информационной технологии – «Квантовым технологиям» (приоритет 1) могут быть выполнены только путем разработки соответствующей системы подготовки инженерных и научных кадров в области квантовых информационных технологий, квантово-сильного искусственного интеллекта и квантовой устойчивости. Создание вертикально интегрированной системы образования, предусматривающей не только получение базового высшего образования, но и подготовку научных кадров и повышение компетенций уже работающих специалистов, будет способствовать развитию инновационных технологий в целом. В статье представлены результаты совместной деятельности научно-технологического университета «Сириус» совместно с Квантовым консорциумом бизнеса и университетов по подготовке кадров в области квантовых информационных технологий для развития такой системы образования.

Ключевые слова: квантовые вычисления, квантовые информационные технологии, информационная безопасность, образовательная программа, новые угрозы квантовой безопасности, квантовая и постквантовая криптография, квантово-сильный искусственный интеллект

Финансирование: Результаты получены при финансовой поддержке проекта «Технологии противодействия ранее неизвестным квантовым киберугрозам», реализуемого в рамках государственной программы федеральной территории «Сириус» «Научно-техническое развитие федеральной территории “Сириус”» (Соглашение № 23-03 от 27 сентября 2024 г.).

Для цитирования: Skiba V.Y., Petrenko S.A., Abakumov E.M. Education system of engineers and scientific personnel in the sphere of quantum information technologies, quantum robust artificial intelligence and quantum resilience // Computing, Telecommunications and Control. 2025. T. 18, № 4. С. 44–52. DOI: 10.18721/JCSTCS.18404

Introduction

Currently, quantum technologies are the driving force behind technological progress (Industry 4.0/5.0/6.0 or Society 6.0). They have already irreversibly changed the world and have spread not only in the scientific community, but also in various fields of human activity, including manufacturing, the military-industrial complex, ecology, medicine and information security [1–11].

The development of quantum information technologies gives rise to new quantum threats to information security [10].

In the Russian Federation, the development of quantum technologies is receiving the closest attention [10, 11]. For instance, in early 2024, the government roadmap for the development of the

high-tech field of Quantum Computing, prepared by the Rosatom State Corporation for 2020, was updated as part of the new federal project "Quantum Technologies as a component of the National Project Data Economics". This federal project replaced the "quantum roadmap", which was set to run until the end of 2024 and was coordinated by Rosatom. The strategic goal of the new policy document is to create a quantum industry by 2030, involving not only scientists and specialists, but also entrepreneurs and a wide range of future consumers of quantum technologies as active participants.

The study [11] examines initiatives to develop a new educational program – "Quantum informatics: Information security". It proposes directing the training of specialists in this field toward two categories of students:

1) developers of core technologies: students will acquire fundamental systems knowledge and a deep understanding of the core technologies in the field, will be capable to develop system components for the technology stack and will be competent enough to contribute to international projects;

2) developers of applied solutions: this category will unite "applied engineers" who will be able to create trusted and secure solutions for various domains (fintech, telecommunications, defense, medical technologies, management, retail, logistics etc.), will have knowledge about existing technology stacks and will be capable to design and develop applied solutions based on them.

The implementation of this initiative in 2024–2025 at the Sirius University of Science and Technology (Sirius University) together with the Quantum Consortium of Enterprises and Universities for training personnel in Quantum Information Technology (Quantum Consortium), has highlighted the need to establish vertically integrated education system for engineers and researchers in quantum information technology, quantum robust artificial intelligence (AI) and quantum resilience of information systems.

Related work and background

According to [11], quantum informatics and information security is a relatively new and rapidly developing field of scientific research that arose at the intersection of quantum mechanics, information theory, programming and information security. Its main branches are quantum computing, quantum communication and quantum information theory.

A bibliometric analysis of academic literature on quantum technologies for 1990–2020 revealed [2, 12] the dynamic growth of the field, high degree of concentration of research and international scientific relations, as well as the involvement of not only universities and academic institutions, but also large corporations (especially from Japan) and military research structures (primarily from the USA). At the same time, Russia exhibits the following characteristics:

- high concentration of research in metropolitan areas and significant international collaboration;
- leading contribution of the Russian Academy of Sciences (RAS), which ranks sixth among scientific organizations in the world in terms of the number of publications in the field of quantum technologies for 1990–2020 [1];
- growing role of universities in the development of the scientific base of quantum and quantum information technologies;
- still limited involvement of the Russian commercial sector in research.

In accordance with the conclusion made in [12], since 2020, the number of scientific publications on the research results in quantum and quantum information technologies has increased exponentially, partly driven by efforts to address information security challenges arising from new quantum threats.

At the same time, the broader landscape is characterized by the absence of mass-production technology for quantum chips and the fact that a dominant physical platform for quantum computers has yet to be established. In parallel, development efforts are underway to create quantum computers

based on more than 10 platforms, the main ones being superconductors, ions, neutral atoms and photons [12].

In Russia, scientific research and engineering studies are also being conducted to create the first domestic quantum computers [3, 4, 10, 12]. Cooperation and consortia are being formed on the basis of domestic centers of competence in quantum technologies, quantum information technologies and quantum communications [12]. Quantum processors with 2–10 qubits and quantum simulators with 10–20 qubits have been developed, and the first domestic quantum processors with 50–100 qubits are expected to appear by the end of 2025.

In order to fulfill the directive of the President of the Russian Federation to create a fundamentally new University of Quantum Technologies dedicated to studying advanced developments in quantum computing and quantum information technologies, as well as to engage school students in the educational process, a new research group "Technologies for Countering Previously Unknown Quantum Cyber Threats" (research group) was established in 2024 at Sirius University under the supervision of prof. S.A. Petrenko (Grand PhD in Engineering).

The main objective of the research group is to create a promising world-class technology to ensure quantum resilience of leading national digital platforms and blockchain ecosystems of the digital economy of Russia, which, unlike existing technologies, will prevent significant or catastrophic consequences in the face of previously unknown cyberattacks carried out by malicious actors using quantum computer [12].

Task statement

From its inception, the research group began to implement the initiatives outlined in [11]. Consequently, in the drafted Concept of ensuring the resilience of operation of national digital platforms and blockchain ecosystems under the new quantum threat to security [12], one of the key tasks is the development of a set of interrelated training programs in quantum information technologies and quantum resilience of national digital platforms and blockchain ecosystems, including supplementary professional education and/or upskilling.

Representatives of the research group participated in the final panel discussion at the 3rd All-Russian Forum "Trusted Quantum Technologies and Communications (Quantum-2025)", dedicated to the specialists training and human resources potential of the industry, on January 30, 2025. The discussion also included representatives of leading universities and training centers – ANO "NTC CC", MISIS, NIO "Quantum Center" MTUSI, TUSUR and the Quantum Consortium, who highlighted the critical shortage of specialists trained in quantum information technologies and quantum robust AI.

In early 2025, Sirius University joined the Quantum Consortium alongside prof. S.V. Ulyanov (Grand PhD in Physics and Mathematics) and A.G. Reshetnikov (PhD in Engineering), the co-authors of a number of works on quantum robust AI and cognitive robotics [8, 9]. This collaboration led to the formulation of the task of developing educational programs spanning from school students to top manager and/or qualified end-users for work in quantum computer science and quantum robust AI.

The jointly defined objective is to establish a vertically integrated education system for engineers and research personnel in quantum information technologies, quantum robust AI and quantum rsilience. This system will range from organizing introductory career guidance events in educational institutions to providing advanced training for existing IT professionals and/or systems engineers in robotics.

In other words, the task entails integrating the educational system (universities) with industry (business) to create a university-based scientific and industrial complex. This complex will serve as the foundation for implementing the innovative "education–science–production" project. At the same time, the vertical integration spans from foundational education systems to specialized retraining centers ans research institutions in active cooperation with production companies implementing quantum information technologies in the product life cycle.

Currently, the task of developing quantum processors or quantum information technologies depends not only on a shortage of engineering personnel with the requisite competencies, but also on the absence of specialists trained to serve as qualified end-users in industry, who could form a request for the development and/or implementation of such technologies in production and their further practical application.

Establishing a vertically integrated education system, which provides not only foundational higher education, but also research training and upskilling for current professionals, enables meeting the requirements of the federal project "Personnel for the Digital Economy or Data Economy" for the key information technology "Quantum Technologies" (priority 1), and will contribute to the development of innovative technologies in general. It is noteworthy that upon acquiring the necessary skills and competencies for solving complex trans-computational problems, the next significant step would be the emergence of quantum AI, that is, AI created on the basis of a quantum computer.

Components of a vertically integrated system and the main stages of its implementation

The first component of the vertically integrated system (Fig. 1) is the implementation of introductory career guidance activities for educational institutions of the "Sirius" federal territory, implementing educational programs of basic general and secondary general education.

The core of this component is the educational and outreach program "New threats to information security using AI technology and quantum computers", which includes the following set of lectures:

- Blockchain ecosystems and their role in the digital economy of Russia;
- Challenges to the resilience of national blockchain ecosystems in the context of emerging quantum security threat;
- Methods and algorithms for the synthesis of quantum-resilient blockchain ecosystems and platforms of the digital economy of Russia;
- Methodology for addressing the synthesis of technologies and programs to ensure quantum resilience of national blockchain ecosystems and platforms of the digital economy of Russia;
- Models for ensuring quantum resilience of national blockchain ecosystems and platforms of the digital economy of Russia.

Holding these events on a regular basis will help engage school students during their studies and attract their interest in pursuing relevant specializations at Sirius University (Fig. 2).

The next component of the vertically integrated system in terms of importance is the training of research and teaching staff, which will enable the recruitment of specialists with higher education in

Fig. 1. Vertically integrated education system

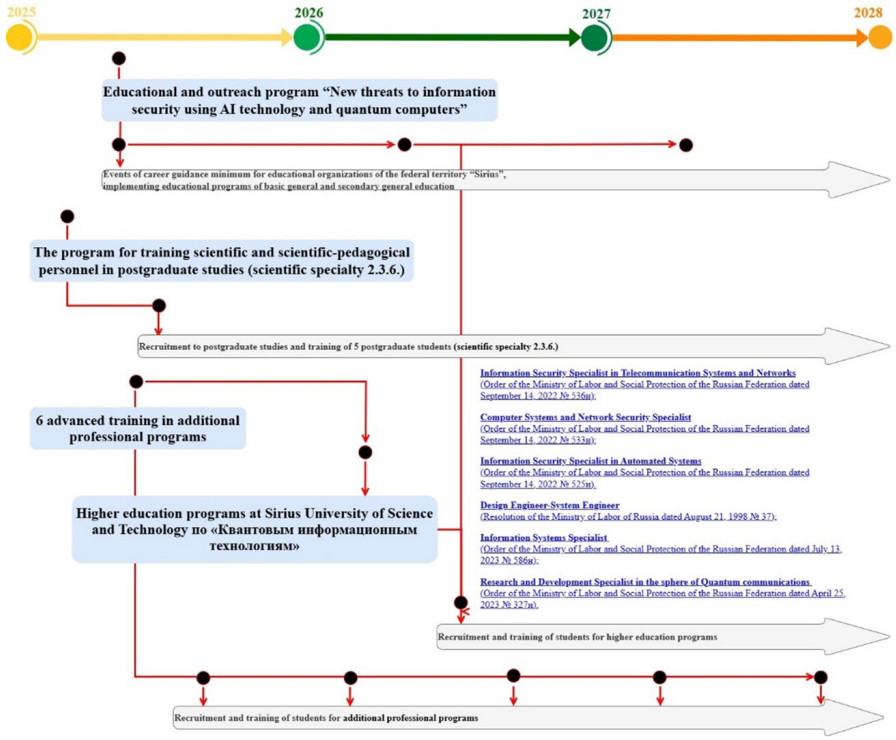


Fig. 2. Relationship between the components of a vertically integrated system and the timing of their implementation

other related IT fields to work in quantum information technologies. Taking into account the importance of ensuring the quantum resilience for digital platforms and blockchain ecosystems, the first program for training research and teaching staff was developed in specialization 2.3.6 "Methods and systems of information protection, information security".

Also of significant importance is accelerated retraining through additional professional programs for professionals and end-users of information technologies. Six additional professional training programs have been developed for the following specialist profiles:

- information security specialist in telecommunication systems and networks (requirements approved by Order No. 536n of the Ministry of Labor and Social Protection of the Russian Federation, dated September 14, 2022);
- computer systems and network security specialist (requirements approved by Order No. 533n of the Ministry of Labor and Social Protection of the Russian Federation, dated September 14, 2022);
- information security specialist in automated systems (requirements approved by Order No. 525n of the Ministry of Labor and Social Protection of the Russian Federation, dated September 14, 2022).
- design engineer/systems engineer (qualification outlined in Decree No. 37 of the Ministry of Labor and Social Protection of the Russian Federation, dated August 21, 1998);
- information systems specialist (requirements approved by Order No. 586n of the Ministry of Labor and Social Protection of the Russian Federation, dated July 13, 2023);
- research and development specialist in quantum communications (requirements approved by Order No. 327n of the Ministry of Labor and Social Protection of the Russian Federation, dated April 25, 2023).

The list of additional professional programs is given in Fig 3.

The final and key component of the vertically integrated system is necessarily a program for training qualified specialists in quantum information technologies to fully and adequately meet the demand for

Fig. 3. List of programs of additional professional education and acquired competencies

experts among end-users and manufacturers. Such training is planned to be carried out in accordance with the higher education program “Quantum Information Technologies”.

Graduates of this program will be qualified specialists in the fields and with the qualifications previously outlined for additional professional training programs.

First results

At a meeting of the Quantum Consortium held at the Russian Union of Industrialists and Entrepreneurs on April 24, 2025, its President, Alexander Shokhin, noted in his welcoming speech that the field of quantum technologies has been developing in Russia for several decades:

“We have formed an expert consensus on key issues and clearly understand that in the coming years, the field of quantum technologies, alongside AI, will be one of the main drivers of development of industry and the Russian economy. There is a lot of work ahead. Together with universities, we need to train teachers and engineers and create favorable conditions for the development of the industry”.

During the meeting, Sirius University presented educational programs for targeted training of personnel in quantum engineering and quantum industrial AI (Fig. 3).

The discussion focused on cultivating demand among industrial companies for highly qualified personnel capable of developing and implementing innovative solutions, including industrial ones, based on end-to-end quantum information technologies.

The participants of the discussion explored the possibilities of training a new generation of engineering personnel to meet the needs of industrial customers on the basis of technical universities in problem-oriented areas, taking into account international practice. Particular attention was paid to the potential for establishing new specialized profiles in quantum information technologies under the umbrella of “Quantum Engineering”.

The presentation of these programs aroused great interest among members of the Quantum Consortium.

In July 2025, the first cohort of postgraduate students was enrolled in specialty 2.3.6 with dissertation research aimed at applying quantum information technologies, quantum algorithms and post-quantum algorithms to address the problem of ensuring the resilience of national digital platforms and blockchain ecosystems.

The enrollment and training of students for the higher education programs are scheduled to commence in the summer of 2026.

Conclusion

The development and most importantly, the implementation of the proposed vertically integrated education system for engineers and research personnel will meet the requirements of the federal project "Personnel for the Digital Economy or Data Economy") concerning key information technology – "Quantum Technologies" (priority 1), and will also contribute to the development of innovative information security technologies and the use of quantum robust AI in industry and robotics.

The postgraduate, master's and additional professional education programs included in the system are aimed primarily at training specialists in developing universal libraries of quantum algorithms regardless of quantum platforms and chips used. Quantum algorithms can be created without deep understanding of quantum physics. It is significant that developers who have the skills to program quantum computers will gain a competitive advantage as "quantum hardware" technology matures.

REFERENCES

1. **Terekhov A.I.** On the Development of the Scientific Base of Quantum Technologies. *Economics of Science*, 2022, Vol. 8, No. 1, Pp. 58–72. DOI: 10.22394/2410-132X-2022-8-1-58-72

2. **Terekhov A.I.** Bibliometric Analysis of Academic Literature on Quantum Information Processing. *Photonics*, 2024, Vol. 18, No. 4, Pp. 296–312. DOI: 10.22184/1993-7296.FRos.2024.18.4.296.312

3. **Holevo A.S.** *Vvedenie v kvantovuiu teoriiu informatsii* [*Introduction to Quantum Information Theory*]. Moscow: MTsNMO, 2002. 128 p.

4. **Holevo A.S.** *Veroiatnostnye i statisticheskie aspekty kvantovoi teorii* [*Probabilistic and statistical aspects of quantum theory*]. Moscow: MTsNMO; NMU, 2020. 364 p.

5. **Bogdanov Y.I., Valiev K.A., Kokin A.A.** Quantum computers: achievements, implementation difficulties, and prospects. *Microelectronics*, 2011, Vol. 40, No. 4, Pp. 243–255. DOI: 10.1134/S1063739711040032

6. **Nielsen M.A., Chang I.L.** *Quantum Computation and Quantum Information: 10th Anniversary ed*. Cambridge: Cambridge University Press, 2011. 702 p.

7. **Bennett C.H., Shor P.W.** Quantum information theory. *IEEE Transactions on Information Theory*, 1998, Vol. 44, No. 6, Pp. 2724–2742. DOI: 10.1109/18.720553

8. **Borovinsky V.V., Kapkov R.Yu., Reshetnikov A.G., Ulyanov S.V.** *Intelligent cognitive robotics*, 2nd ed. Moscow: KURS, 2024. 557 p.

9. **Dobrynin V., Reshetnikov G., Sacharov Yu., Ulyanov S.** Quantum and soft computational intelligence technologies in intelligent control processes design. *Systems analysis in science and education*, 2010, No. 2, Pp. 1–19.

10. **Skiba V.Yu., Petrenko S.A., Murzina A.A., Popova K.R.** New types of threats and assessment of quantum stability of information systems in the field of foreign trade activity. *Computing, Telecommunications and Control*, 2024, Vol. 17, No. 4, Pp. 16–34. DOI: 10.18721/JCSTCS.17402

11. **Petrenko S.A., Petrenko A.S., Ozhiganova M.I.** On the development of a new educational program "Quantum information science: information security". *Zaŝita informacii. Inside*, 2024, Vol. 118, No. 4, Pp. 17–23.

12. **Skiba V.Yu., Petrenko S.A., Gnidko K.O., Petrenko A.S.** Concept of ensuring the resilience of operation of national digital platforms and blockchain ecosystems under the new quantum threat to security. *Computing, Telecommunications and Control*, 2025, Vol. 18, No. 2, Pp. 56–73. DOI: 10.18721/JCSTCS.18205

INFORMATION ABOUT AUTHORS / СВЕДЕНИЯ ОБ АВТОРАХ

Vladimir Yu. Skiba
Скиба Владимир Юрьевич
E-mail: vskiba69@mail.ru
ORCID: https://orcid.org/0000-0002-9805-7800

Sergei A. Petrenko
Петренко Сергей Анатольевич
E-mail: petrenko.sa@talantiuspeh.ru
ORCID: https://orcid.org/0000-0003-0644-1731

Evgeny M. Abakumov
Абакумов Евгений Михайлович
E-mail: abakumov.em@talantiuspeh.ru

Submitted: 09.09.2025; Approved: 29.10.2025; Accepted: 14.11.2025.
Поступила: 09.09.2025; Одобрена: 29.10.2025; Принята: 14.11.2025.

Computing, Telecommunication and Control, 2025, Vol. 18, No. 4, Pp. 53–66.
Информатика, телекоммуникации и управление. 2025. Том 18, № 4. С. 53–66.

Circuits and Systems for Receiving, Transmitting and Signal Processing

Устройства и системы передачи, приема и обработки сигналов

Research article
DOI: https://doi.org/10.18721/JCSTCS.18405
UDC 621.391.8

LOW COMPUTATIONAL COMPLEXITY TECHNIQUE BASED ON A POLYPHASE STRUCTURE FOR MODULATION AND DEMODULATION OF FBMC/OQAM-OTFS SIGNALS

B.T. Khuc ✉ , *A.L. Gelgor*

St. Petersburg State Polytechnical University,
St. Petersburg, Russian Federation

✉ khucbang@mail.ru

Abstract. The paper proposes a low computational complexity technique based on a polyphase structure for modulation and demodulation of FBMC/OQAM-OTFS signals. This approach effectively reduces the overall computational complexity when compared to the frequency spreading approach (FS-FBMC/OQAM-OTFS), which in turn outperforms the direct modulation and demodulation approach for FBMC/OQAM-OTFS signals (Direct-FBMC/OQAM-OTFS). Simulation results explicitly demonstrate that, for an overlapping factor of $K = 4$, various PPN-FBMC/OQAM variants can indeed achieve a 2.5–4 times reduction in the computational complexity with an energy loss of no more than 1 dB compared to FS-FBMC/OQAM-OTFS. These obtained results are observed under standard multipath channel profiles such as EPA, EVA, and ETU in both moderately and highly dynamic scenarios. These findings suggest that the PPN-FBMC/OQAM-OTFS technique is a feasible and promising alternative to conventional OFDM in high-mobility wireless scenarios.

Keywords: OFDM, OTFS, FS-FBMC/OQAM-OTFS, PPN-FBMC/OQAM-OTFS, highly dynamic channel

Citation: Khuc B.T., Gelgor A.L. Low computational complexity technique based on a polyphase structure for modulation and demodulation of FBMC/OQAM-OTFS signals. Computing, Telecommunications and Control, 2025, Vol. 18, No. 4, Pp. 53–66. DOI: 10.18721/JCSTCS.18405

Научная статья
DOI: https://doi.org/10.18721/JCSTCS.18405
УДК 621.391.8

ВЫЧИСЛИТЕЛЬНО ЭФФЕКТИВНЫЙ АЛГОРИТМ НА ОСНОВЕ ПОЛИФАЗНОЙ СТРУКТУРЫ ДЛЯ МОДУЛЯЦИИ И ДЕМОДУЛЯЦИИ СИГНАЛОВ FBMC/OQAM-OTFS

Б.Т. Хук ✉ , *А.Л. Гельгор*

Санкт Петербургский государственный политехнический университет,
Санкт-Петербург, Российская Федерация

✉ khucbang@mail.ru

Аннотация. В данной статье предлагается новый и перспективный подход на основе полифазной структуры для вычислительно эффективной модуляции и демодуляции сигналов FBMC/OQAM-OTFS, называемый PPN-FBMC/OQAM-OTFS. Предлагаемый подход обеспечивает снижение вычислительной сложности по сравнению с подходом на основе расширения частоты FS-FBMC/OQAM-OTFS, который, в свою очередь, уже превосходит метод прямой модуляции и демодуляции сигналов FBMC/OQAM-OTFS. Результаты моделирования проведенных исследований наглядно показывают, что при использовании прототипа-фильтра Phydyas с коэффициентом перекрытия K = 4 различные варианты PPN-FBMC/OQAM-OTFS действительно могут обеспечить существенный выигрыш в вычислительной сложности в 2,5–4 раза при энергетических потерях не более 1 дБ по сравнению с FS-FBMC/OQAM-OTFS. Эти результаты были получены при стандартных многолучевых каналах, таких как EPA, EVA и ETU как в умеренно-, так в высоко-динамичных сценариях. Полученные результаты позволяют предположить, что технология PPN-FBMC/OQAM-OTFS является работоспособной и перспективной альтернативой традиционному OFDM в условиях беспроводной связи с высокой мобильностью.

Ключевые слова: OFDM, OTFS, FS-FBMC/OQAM-OTFS, PPN-FBMC/OQAM-OTFS, высоко-динамичный канал

Для цитирования: Khuc B.T., Gelgor A.L. Low computational complexity technique based on a polyphase structure for modulation and demodulation of FBMC/OQAM-OTFS signals // Computing, Telecommunications and Control. 2025. T. 18, № 4. C. 53–66. DOI: 10.18721/JCSTCS.18405

Introduction

Currently, more diverse transmission channel conditions have emerged in communication systems including highly dynamic channels due to increased user speeds [1]. This causes the transmission channel to be frequently affected by time and frequency selectivity, resulting from multipath and Doppler effects [2].

In modern communication systems, including 5G New Radio, Orthogonal Frequency Division Multiplexing (OFDM) modulation remains the preferred choice due to its ability to mitigate Inter-Symbol Interference (ISI) and its simplicity in design [3]. However, OFDM modulation suffers severe performance degradation in highly dynamic channels caused by Inter-Carrier Interference (ICI). Additionally, OFDM provides high Out of Band Emission (OOBE) levels due to the use of rectangular pulse for the signal generation and low spectral efficiency due to the use of cyclic prefix [3, 4]. Therefore, in recent works [5, 6], a new modulation technique called FBMC/OQAM-OTFS (Filter-Bank Multi-Carrier with Offset Quadrature Amplitude Modulation and Orthogonal Time-Frequency Space pre-processing) has been proposed to overcome these limitations. This technique has the following advantages:

1) it provides low OOBE levels due to the use of prototype filter for each subcarrier;

2) cyclic prefixes are not required, leading to an increased spectral efficiency;

3) its Bit Error Rate (BER) performance doesn't fall, i.e., doesn't saturate, even in highly dynamic wireless channels.

Despite its significant performance advantages, the direct implementation of FBMC/OQAM-OTFS (Direct-FBMC/OQAM-OTFS) faces major limitations in terms of computational complexity. Specifically, it requires extremely high computational demands, as well as substantial processing and latency burdens for the system [6, 7]. The work [7] has shown that the computational cost of the Direct-FBMC/OQAM-OTFS modulation/demodulation scheme can be up to 10^4 times higher than that of conventional OFDM schemes. Therefore, the implementation of Direct-FBMC/OQAM-OTFS in next-generation communication systems is practically infeasible.

A commonly adopted approach for implementing the FBMC/OQAM-OTFS technique is based on frequency spreading (FS) FBMC/OQAM-OTFS. This approach aims to filter the real and imaginary parts of information symbols in the frequency domain by upsampling and filtering before the Inverse Fast Fourier Transform (IFFT) operation with the extended size. The FS-FBMC/OQAM-OTFS allows a reduction of the computational complexity compared to the direct method by up to hundreds of times [7, 8]. However, the computational complexity of the FS-FBMC/OQAM-OTFS technique is still very high compared to the OFDM modulation (about 11–16 times higher) due to the use of the Fast Fourier Transform (FFT)/IFFT blocks of extended-size [4]. As a result, the adaptation of FBMC/OQAM-OTFS in next-generation mobile communication systems continues to face considerable challenges.

In this work, we propose two variants of the FBMC/OQAM-OTFS implementation based on a polyphase network structure (PPN-FBMC/OQAM-OTFS). Unlike the FS-FBMC/OQAM-OTFS approach, in the PPN-FBMC/OQAM-OTFS approach, the filtering process is implemented by the filter bank based on the polyphase structure, similar to how it is done for FBMC/OQAM [7]. The proposed PPN-FBMC/OQAM-OTFS schemes achieve a reduction in computational complexity by a factor of 2.5 to 4, with less than 1 dB energy loss across various multipath channel profiles, compared to FS-FBMC/OQAM-OTFS.

Direct form of FBMC/OQAM-OTFS implementation

Through this paper, we assume that one FBMC/OQAM-OTFS framework has a bandwidth B_{W} and a period T_{Frame}, in which B_{W} contains M subcarriers and T_{Frame} is divided into N sub-symbols (i.e., $B_{\mathrm{W}} = M\Delta f$, $T_{\mathrm{Frame}} = NT$). Then, the Doppler and delay steps can be expressed as $1/NT$ and $1/M\Delta f$, respectively. The transformation chain in system using FBMC/OQAM-OTFS is shown in Fig. 1. Fig. 2 shows the Delay-Doppler (DD) grid that is used to map/demap the information symbols in the OTFS Pre/Post-Processing procedure.

In the FBMC/OQAM-OTFS technique, QAM information symbols are first mapped into the DD domain – $x^{\mathrm{DD}}[n, l]$, where $n = 0, ..., N - 1$, $l = 0, ..., L - 1$. Then, $x^{\mathrm{DD}}[n, l]$ are converted in the time-frequency (TF) domain $X^{\mathrm{TF}}[m, k]$ by using inverse symplectic FFT (ISFFT) [1]:

$$X^{\mathrm{TF}}[l,k] = \frac{1}{\sqrt{NL}} \sum_{n=0}^{N-1} \sum_{l=0}^{L-1} x^{\mathrm{DD}}[n,l] \exp\left(-j2\pi \left(\frac{nm}{N} - \frac{lk}{K} \right) \right). \quad (1)$$

In an FBMC/OQAM system, OQAM pre-processing is performed in the TF domain by shifting the real and imaginary parts of QAM symbols by half of the symbol period, $T/2$. As a result, after OQAM pre-processing, the symbol count is doubled compared to a conventional OFDM modulation, and the symbol period becomes equal to $T/2$ [5, 6]:

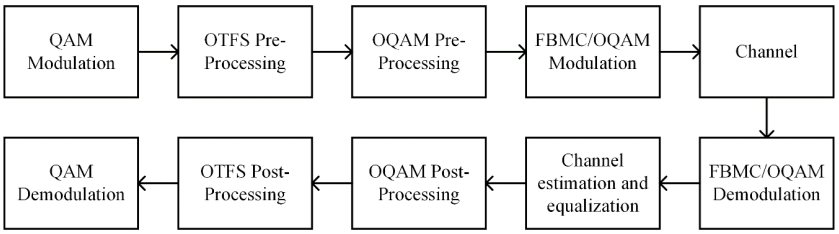


Fig. 1. FBMC/OQAM-OTFS scheme

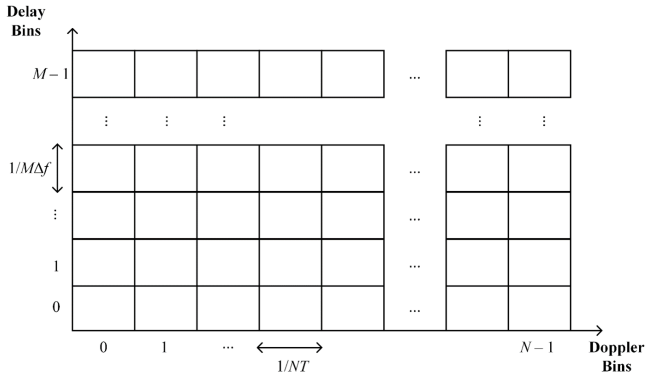


Fig. 2. OTFS DD grid

$$\begin{aligned}&\left\{X^{\mathrm{TF}}\left[m,k\right],\ m=0,\ \ldots,\ N-1,\ k=0,\ \ldots,\ L-1\right\}\rightarrow\\&\rightarrow\left\{X^{\mathrm{TF}}\left[m,k'\right],\ m=0,\ \ldots,\ N-1,\ k'=0,\ \ldots,\ 2K-1\right\}.\end{aligned}\tag{2}$$

Finally, the transmitted FBMC/OQAM-OTFS signal is generated by using the Heisenberg transformation [5, 6]:

$$s\left(t\right)=\sum_{m=0}^{N-1}\sum_{k'=0}^{2L-1}X^{\mathrm{TF}}\left[m,k'\right]g^{\mathrm{TX}}\left(t-\frac{k'T}{2}\right)\exp\left(j2\pi m\Delta f\left(t-\frac{k'T}{2}\right)\right),\tag{3}$$

where $g^{\mathrm{TX}}(t)$ is the transmit prototype filter.

At the receiver, the time-domain signal without a noise component can be represented as follows [5, 6]:

$$r\left(t\right)=\int\int h\left(\tau,\nu\right)\exp\left(j2\pi\nu\left(t-\tau\right)\right)s\left(t-\tau\right)d\tau d\nu,\tag{4}$$

where $h(\tau, \nu)$ is the channel response in the DD domain which in turn can be represented as follows [5, 6]:

$$h\left(\tau,\nu\right)=\sum_{p=1}^{P}h_p\delta\left(t-\tau_p\right)\delta\left(\nu-\nu_p\right),\tag{5}$$

where h_p, τ_p, and ν_p are the average path gain, path delay and path Doppler shift of the p-th path; and P is the number of signal propagation paths.

To detect information symbols in the TF domain, the matched filter and Wigner transformation are used [5, 6]:

$$G(\tau,\nu)=\int \exp\left(-j2\pi(t-\tau)\right)g^{\mathrm{RX}}(t-\tau)r(t)\,dt; \tag{6}$$

$$Y^{\mathrm{TF}}[l,k']=G(\tau,\nu)\big|_{\tau=k'T/2,\ \nu=m\Delta f}. \tag{7}$$

By applying OQAM post-processing, the QAM information symbols are taken as follows [5, 6]

$$\begin{aligned}&\left\{Y^{\mathrm{TF}}[m,k'],\ m=0,\ \ldots,\ N-1,\ k'=0,\ \ldots,\ 2L-1\right\}\rightarrow\\&\rightarrow\left\{Y^{\mathrm{TF}}[m,k],\ m=0,\ \ldots,\ N-1,\ k=0,\ \ldots,\ L-1\right\}.\end{aligned} \tag{8}$$

Then the received information symbols in the DD domain are obtained by using symplectic FFT (SFFT) as follows [5, 6]:

$$y^{\mathrm{DD}}[n,l]=\frac{1}{\sqrt{NL}}\sum_{n=0}^{N-1}\sum_{l=0}^{L-1}Y^{\mathrm{TF}}[m,k]\exp\left(-j2\pi\left(\frac{nm}{N}-\frac{lk}{L}\right)\right). \tag{9}$$

The Direct-FBMC/OQAM-OTFS modulation/demodulation incurs extremely high complexity due to the need for 2D transform calculations in [1, 9]. Therefore, this issue needs to be overcome before FBMC/OQAM-OTFS can be practically deployed in next-generation mobile networks.

In FBMC/OQAM-OTFS technique, the OTFS 2D-transformations add only a moderate computational load, whereas the FBMC/OQAM, which is a combination of 2D-transformations with the extended size and prototype filtering, dominates the overall complexity. Therefore, the focus here is on simplifying the FBMC/OQAM scheme.

FBMC/OQAM-OTFS implementation with frequency spreading

An alternative approach for implementing the FBMC/OQAM-OTFS modulation/demodulation is the frequency spreading technique (FS-FBMC/OQAM-OTFS) [7]. In this approach, the real and imaginary parts of the information symbols are processed in the frequency domain through an upsampling and filtering operation performed prior to the IFFT with an extended size. The structure of the FS-FBMC/OQAM-OTFS scheme is illustrated in Fig. 3.

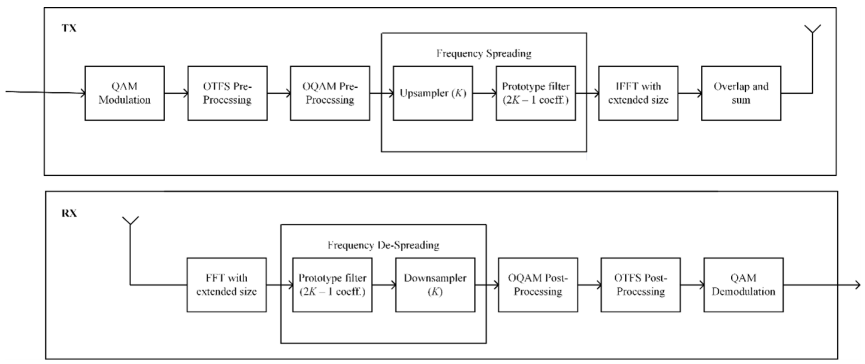


Fig. 3. FS-FBMC/OQAM-OTFS scheme

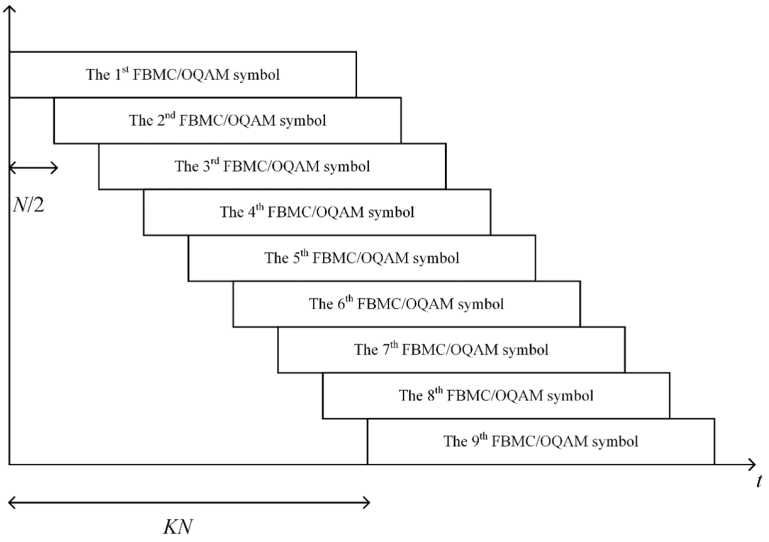


Fig. 4. Illustration of summing of overlaping FBMC/OQAM-OTFS symbols

At the transmitter, at the first stage, the OQAM pre-processing is used after ISFFT operation, where the in-phase and quadrature components of the QAM symbols are time-shifted by half the symbol duration. As a result, the orthogonality between subcarriers is preserved. After OQAM pre-processing, each OQAM symbol is upsampled by the overlapping factor of K and filtered by a prototype filter in the frequency domain. This combination of upsampling and filtering process constitutes the frequency spreading operation. Accordingly, the IFFT size is extended by a factor of K. As can be seen from Fig. 4, at the final stage, the transmitted signal is formed by summing of overlapping FS-FBMC/OQAM-OTFS symbols following each other with a step of $N/2$ samples. At the receiver, the received signal is processed in the reverse order of the transmitter operations, with the real and imaginary parts being handled separately.

Proposed low computational complexity FBMC/OQAM-OTFS implementation

The FS-FBMC/OQAM-OTFS technique is still ineffective in terms of the computational complexity due to the use of frequency spreading and FFT/IFFT operations with the size KN. Therefore, in this work, we propose two alternative implementation approaches for the FBMC/OQAM-OTFS technique based on a polyphase structure (PPN-FBMC/OQAM-OTFS).

The filters used for FBMC/OQAM-OTFS modulation are a finite impulse response (FIR) filters with the length $P = KN$. The relationship between the input and output of the filter is expressed as follows:

$$y[n] = \sum_{i=0}^{P-1} h_i x[n]. \tag{10}$$

Applying the Z-transformation to (10), we obtain

$$\begin{aligned} H(Z) &= \sum_{i=0}^{P-1} h[i] Z^{-i} = \sum_{n=0}^{N-1} \sum_{k=0}^{K-1} h[kN+n] Z^{-(kN+n)} = \\ &= \sum_{n=0}^{N-1} \left(\sum_{k=0}^{K-1} h[kN+n] Z^{-kN} \right) Z^{-n} = \sum_{n=0}^{N-1} E_n\left(Z^N\right) Z^{-n}, \end{aligned} \tag{11}$$

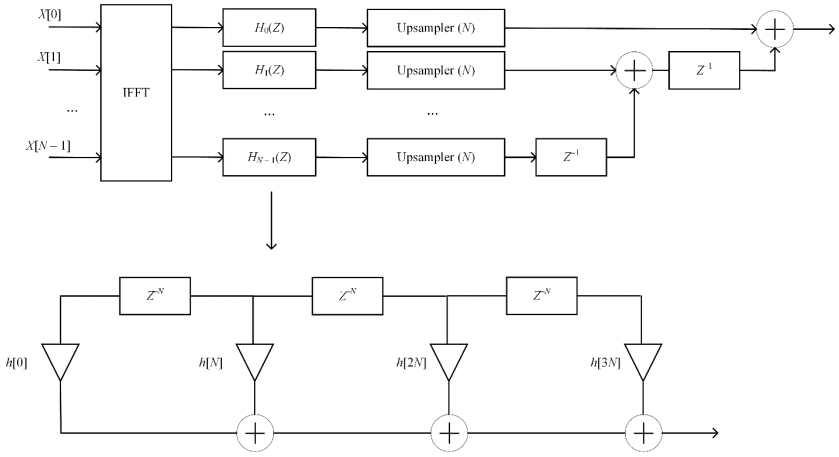


Fig. 5. Polyphase structure in PPN-FBMC/OQAM-OTFS transmitter

where $E_n\left(Z^N\right)=\sum_{k=0}^{K-1}h\left(kN+n\right)Z^{-kN}$ is the elementary component of the filter $H(Z)$. As seen from (11), the filter $H(Z)$ can be decomposed into N elementary components, as depicted in Fig. 5.

Consider a version of the $H_i(Z)$ filter, which is obtained by shifting of $H(Z)$ by i/N in the frequency domain. Using the polyphase representation, $H_i(Z)$ filter is defined as:

$$H_i\left(Z\right)=\sum_{i=0}^{P-1}h\left[i\right]e^{j2\pi i/N}Z^{-i}=\sum_{n=0}^{N-1}\sum_{k=0}^{K-1}h\left[kN+n\right]e^{j2\pi i\left(kN+n\right)/N}Z^{-\left(kN+n\right)}=$$
$$=\sum_{n=0}^{N-1}e^{j2\pi ni/N}E_n\left(Z^N\right)Z^{-n}. \tag{12}$$

Denote $W=\exp(-j2\pi/N)$. From (12), it follows that the filter bank is obtained by shifting $H(Z)$ in the frequency domain by a multiple of $1/N$ and is defined as follows:

$$\begin{bmatrix}H_0\left(Z\right)\\H_1\left(Z\right)\\\cdots\\H_{N-1}\left(Z\right)\end{bmatrix}=\begin{bmatrix}1&1&\cdots&1\\1&W^{-1}&\cdots&W^{-\left(N-1\right)}\\\cdots&&&\\1&W^{-\left(N-1\right)}&\cdots&W^{-\left(N-1\right)^2}\end{bmatrix}\begin{bmatrix}E_0\left(Z^N\right)\\Z^{-1}E_1\left(Z^N\right)\\\cdots\\Z^{-\left(N-1\right)}E_{N-1}\left(Z^N\right)\end{bmatrix}. \tag{13}$$

Obviously, matrix **W** is the matrix of an IFFT operation. Accordingly, the corresponding filter bank structure at the transmitter side is illustrated in Fig. 6. At the receiver side, the filter bank corresponds to an FFT operation as described in (14), and its structure is also depicted in Fig. 6.

$$\begin{bmatrix}H_0\left(Z\right)\\H_{-1}\left(Z\right)\\\cdots\\H_{-\left(N-1\right)}\left(Z\right)\end{bmatrix}=\begin{bmatrix}1&1&\cdots&1\\1&W&\cdots&W^{N-1}\\\cdots&&&\\1&W^{N-1}&\cdots&W^{\left(N-1\right)^2}\end{bmatrix}\begin{bmatrix}E_0\left(Z^N\right)\\Z^{-1}E_1\left(Z^N\right)\\\cdots\\Z^{-\left(N-1\right)}E_{N-1}\left(Z^N\right)\end{bmatrix}. \tag{14}$$

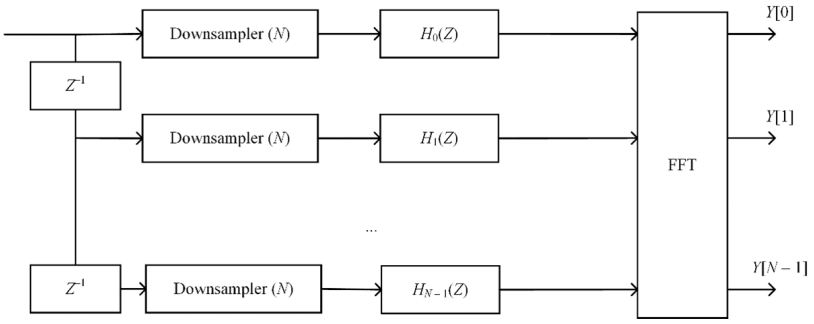


Fig. 6. Polyphase structure in PPN-FBMC/OQAM-OTFS receiver

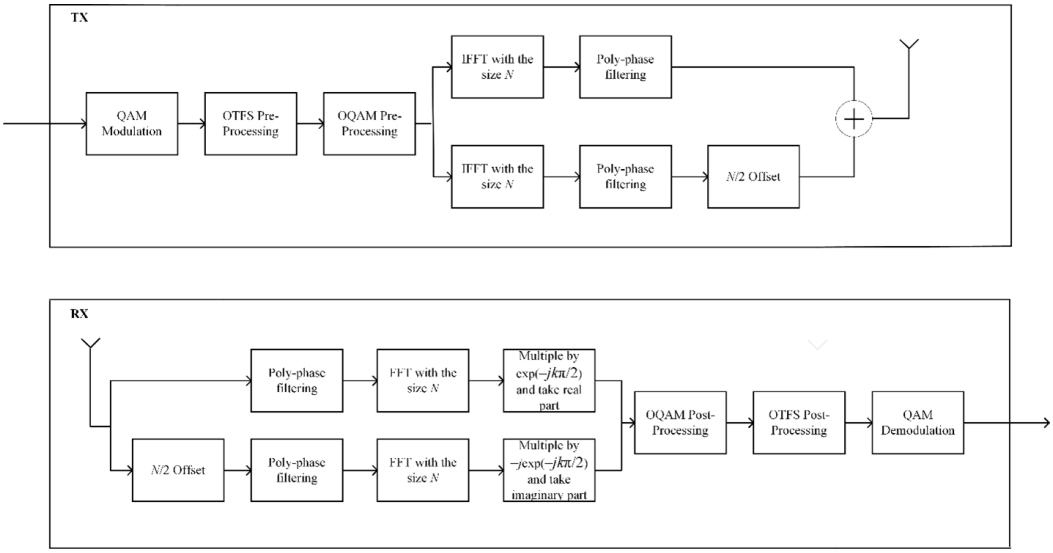


Fig. 7. PPN-FBMC/OQAM-OTFS scheme

Unlike the approach based on the frequency spreading, the PPN-FBMC/OQAM-OTFS technique uses only two FFT/IFFT operations with the size of N, thus significantly reducing computational complexity of the implementing the FBMC/OQAM-OTFS. The scheme of the PPN-FBMC/OQAM-OTFS technique is shown in Fig. 7.

This work also proposes an approach to further reduce the computational complexity of the PPN-FBMC/OQAM-OTFS transmitter, referred to as the low-complexity PPN-FBMC/OQAM-OTFS scheme. In this approach, complex information symbols are used instead of two OQAM symbols as the input of IFFT block, which reduces the cost of two N-IFFT operations to one N-IFFT operation with additional signal processing. To achieve this, the principle of computing the discrete inverse Fourier transforms of two real functions simultaneously using a single IFFT block, as described in [9], is utilized. The multiplication with the phase rotation from OQAM pre-processing in the frequency domain can be replaced by a circular shift of $N/4$ in time domain [10]. The scheme of low-complexity PPN-FBMC/OQAM-OTFS transmitter is shown in Fig. 8.

Computational complexity comparison

An important factor in the computational complexity analysis is the agreement about computational complexity of the multiplication of two complex numbers. Usually, it is used two common ways: either

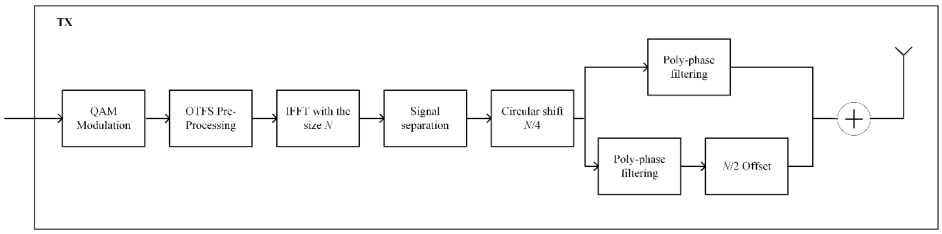


Fig. 8. Low-complexity PPN-FBMC/OQAM-OTFS transmitter

using three additions and three multiplications or using two additions and four multiplications of real numbers [7]. Since the cost of multiplication is significantly higher than that of addition [7], this work adopts the first method. Based on this assumption, the computational complexity of different considered modulation/demodulation techniques is summarized in Tables 1 and 2.

Table 1

Computational complexity comparison on the transmitter side

	Number of real multiplications	**Number of real additions**
OFDM Tx	$L(N(\log_2 N - 3) + 4)$	$L(3N(\log_2 N - 1) + 4)$
OTFS Tx	$3L(N(\log_2 N - 3) + 4)$	$3L(3N(\log_2 N - 1) + 4)$
Direct-FBMC/OQAM Tx	$6LN^2K^2 + 4LN(K - 1)$	$6LN^2K^2 + N(2K - 1)(2L - 1)$
FS-FBMC/OQAM Tx	$2L(NK(\log_2 NK - 3) + + 4 + 2N(K - 1))$	$2L(3NK(\log_2 NK - 1) + 4) + + 2N(K - 1)(2L - 1)$
PPN-FBMC/OQAM Tx	$2L(2NK + N(\log_2 N - 3) + 4)$	$2L(3N(\log_2 N - 1) + 4 + 2N(K - 1) + + 2N(L + K - 2)$
Low-complexity PPN-FBMC/OQAM Tx	$L(N(\log_2 N - 3) + 4 + 4NK)$	$L(3N(\log_2 N - 1) + 4 + 4N(K - 1) + + 4N) + 2N(L + K - 2)$
Direct-FBMC/OQAM-OTFS Tx	$6LN^2(1 + K^2) + 4LN(K - 1)$	$6LN^2(1 + K^2) + N(2K - 1)(2L - 1)$
FS-FBMC/OQAM-OTFS Tx	$2L(N(\log_2 N - 3) + 4) + +2L(NK(\log_2 NK - 3) + + 4 + 2N(K - 1))$	$2L(3N(\log_2 N - 1) + 4) + + 2L(3NK(\log_2 NK - 1) + 4) + + 2N(K - 1)(2L - 1)$
PPN-FBMC/OQAM-OTFS Tx	$2L(N(\log_2 N - 3) + 4) + + 2L(2NK + (N(\log_2 N - 3) + 4))$	$2L(3N(\log_2 N - 1) + 4) + 2L(3N(\log_2 N - 1) + + 4 + 2N(K - 1)) + 2N(L + K - 2)$
Low-complexity PPN-FBMC/OQAM-OTFS Tx	$2L(N(\log_2 N - 3) + 4) + L(N(\log_2 N - 3) + 4 + 4NK)$	$2L(3N(\log_2 N - 1) + 4) + L(3L(\log_2 N - 1) + 4 + 4N(K - 1) + 4N) + 2N(L + K - 2)$

Tables 1 and 2 show the dependences of the number of required operations on the values of N_{FFT} (in Tables 1 and 2 it is marked as N to abbreviate notation), L and K. It is clear that the FS and PPN approaches allow us to reduce the required costs, but it is difficult to draw an exact conclusion. Therefore, Table 3 provides a comparison of the computational complexity of various modulation techniques in terms of the number of real multiplications and additions for $N_{\text{FFT}} = 256$, $L = 20$, and $K = 4$. The results indicate that the computational complexity of OTFS modulation/demodulation is three times higher than that of OFDM for both real multiplications and additions. Regarding FBMC/OQAM techniques, the direct methods (Direct-FBMC/OQAM and Direct-FBMC/OQAM-OTFS) exhibit extremely high computational costs, requiring 4902.3 and 5208.6 times more real multiplications and 1169.4 and 1242.5 times more real additions, respectively, compared to OFDM. By contrast, the FS-FBMC/OQAM and FS-FBMC/OQAM-OTFS schemes substantially reduce computational complexity relative to the direct methods but still demand 13.6–15.6 times more real multiplications and

10.9–12.9 times more real additions than for OFDM, remaining considerably higher number of computations than both OFDM and OTFS. Notably, approaches based on the polyphase network (PPN-FBMC/OQAM and PPN-FBMC/OQAM-OTFS), along with their simplified versions, achieve significant reductions in computational complexity, requiring only 1.4–7.2 times more operations than for OFDM. In particular, the low-complexity PPN-FBMC/OQAM-OTFS scheme requires only 6.2 times more real multiplications and 3.4 times more real additions than OFDM.

Table 2

Computational complexity comparison on the receiver side

	Number of real multiplications	**Number of real additions**
OFDM Rx	$L(N(\log_2 N - 3) + 4)$	$L(3N(\log_2 N - 1) + 4)$
OTFS Rx	$3L(N(\log_2 N - 3) + 4)$	$3L(3N(\log_2 N - 1) + 4)$
Direct-FBMC/OQAM Rx	$6LN^2K^2 + 4LN(K - 1)$	$6LN^2K^2$
FS-FBMC/OQAM Rx	$2L(NK(\log_2 NK - 3) + 4 + 2N(K - 1))$	$2L(3NK(\log_2 NK - 1) + 4) + + 2N(K - 1)(2L - 1)$
PPN-FBMC/OQAM Rx	$2L(2NK + N(\log_2 N - 3) + 4)$	$2L(3N(\log_2 N - 1) + 4 + 2N(K - 1) + + 2N(L + K - 2))$
Low-complexity PPN-FBMC/OQAM Rx	$L(N(\log_2 N - 3) + 4 + 4NK)$	$L(3N(\log_2 N - 1) + 4 + 2N(K - 1) + + 2N(L + K - 2))$
Direct-FBMC/OQAM-OTFS Rx	$6LN^2(1 + K^2) + 4LN(K - 1)$	$6LN^2(1 + K^2)$
FS-FBMC/OQAM-OTFS Rx	$2L(N(\log_2 N - 3) + 4) + + 2L(NK(\log_2 NK - 3) + 4 + + 2N(K - 1))$	$2L(3N(\log_2 N - 1) + 4) + +2L(3NK(\log_2 NK - 1) + 4) + 2N(K - 1)(2L - 1)$
PPN-FBMC/OQAM-OTFS Rx	$2L(N(\log_2 N - 3) + 4) + 2L(2NK + + (N(\log_2 N - 3) + 4))$	$2L(3N(\log_2 N - 1) + 4) + 2L(3N(\log_2 N - 1) + + 4 + 2N(K - 1)) + 2N(L + K - 2)$
Low-complexity PPN-FBMC/OQAM-OTFS Rx	$2L(N(\log_2 N - 3) + 4) + + L(N(\log_2 N - 3) + 4 + 4NK)$	$2L(3N(\log_2 N - 1) + 4) + L(3N(\log_2 N - 1) + + 4 + 2N(K - 1)) + 2N(L + K - 2)$

Table 3

Computational complexity comparison for the case of $N = 256$, $L = 20$, and $K = 4$

	Number of real multiplications divided by such a value for OFDM	**Number of real additions divided by such a value for OFDM**
OFDM	1	1
OTFS	3	3
Direct-FBMC/OQAM	4902.3	1169.4
FS-FBMC/OQAM	13.6	10.9
PPN-FBMC/OQAM	5.2	2.6
Low-complexity PPN-FBMC/OQAM	4.2	1.4
Direct FBMC/OQAM-OTFS	5208.6	1242.5
FS-FBMC/OQAM-OTFS	15.6	12.9
PPN-FBMC/OQAM-OTFS	7.2	4.6
Low-complexity PPN-FBMC/OQAM-OTFS	6.2	3.4

Simulation results and discussion

The simulation parameters are summarized in Table 4.

Table 4

Simulation parameters

Parameters	Value
Modulation technique	CP-OFDM, OTFS, FBMC/OQAM, Direct-FBMC/OQAM-OTFS, FS-FBMC/OQAM-OTFS, PPN-FBMC/OQAM-OTFS, low-complexity FBMC/OQAM-OTFS
N_{FFT}	256
Number of active subcarriers, M	200
Number of symbols, N	20
Length of cyclic prefix in samples (for CP-OFDM)	20
Modulation	QPSK
Equalization method	LMMSE
FEC	LDPC with code rate 1/2
Prototype filter	Phydyas
FBMC/OQAM overlapping factor, K	4
Channel model	EPA/EVA/ETU
Maximal Doppler shift (Hz)	5 – 3000

Shapes of signal spectrum

Spectrum shapes of different modulations are shown in Fig. 9. According to it, all combinations of modulations are divided into two groups: containing and not containing FBMC. That is the first group includes OFDM, OTFS signals, and the second group includes Direct FBMC/OQAM, Direct-FBMC/OQAM-OTFS, FS-FBMC/OQAM-OTFS, PPN-FBMC/OQAM-OTFS and low-complexity PPN-FBMC/OQAM-OTFS signals. Spectrums of each group are almost the same. Modulations containing FBMC provide up to about 100 dB lover level of OOBE with respect to others. Further reduction of the OOBE level can be achieved by applying a higher value of K.

BER Performance

Fig. 10 shows the BER performance of OFDM, OTFS and FBMC/OQAM-OTFS signals under the EVA channel model at Doppler shifts of 70 and 700 Hz, assuming perfect channel estimation. As illustrated, the Direct-FBMC/OQAM-OTFS, FS-FBMC/OQAM-OTFS and OTFS signals exhibit identical BER performance, while the PPN-FBMC/OQAM-OTFS and low-complexity PPN-FBMC/OQAM-OTFS signals likewise achieve the same BER performance. Direct-FBMC/OQAM-OTFS, FS-FBMC/OQAM-OTFS and OTFS signals provide energy gain with respect to CP-OFDM signal, about 1.52 and 1.35 dB at the BER = 10^{-4} for the Doppler shifts of 70 and 700 Hz, respectively. Meanwhile, using PPN-FBMC/OQAM-OTFS and low-complexity PPN-FBMC/OQAM-OTFS signals provides an energy loss of only about 0.81 and 0.67 dB compared to the OTFS, Direct-FBMC/OQAM-OTFS and FS-FBMC/OQAM-OTFS signals for the Doppler shifts of 70 and 700 Hz, respectively. It can be explained that the frequency spreading implementation allows for a more accurate control of the filter bank's frequency response by manipulating each subcarrier individually in the frequency domain. This leads to improvement suppression of ICI compared to the conventional PPN implementation[1] [11].

[1] Prototype filter and structure optimization, Available: http://www.ict-phydyas.org/delivrables/PHYDYAS-D5.1.pdf (Accessed 04.12.2025)

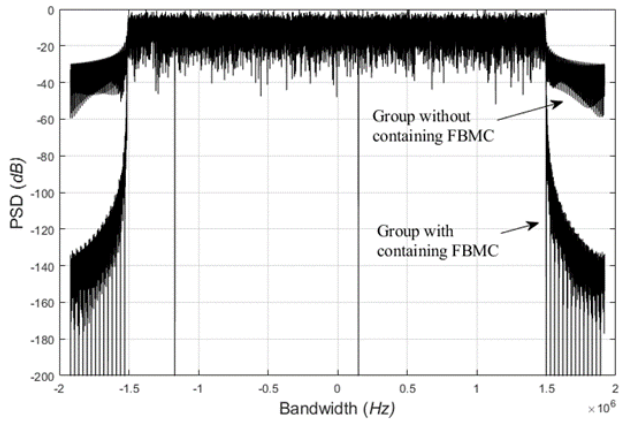


Fig. 9. Spectrum shapes of OFDM, OTFS, FBMC and their combinations. $N_{FFT} = 256$ and the number of active subcarriers $M = 200$. For FBMC Phydyas filter is used and $K = 4$

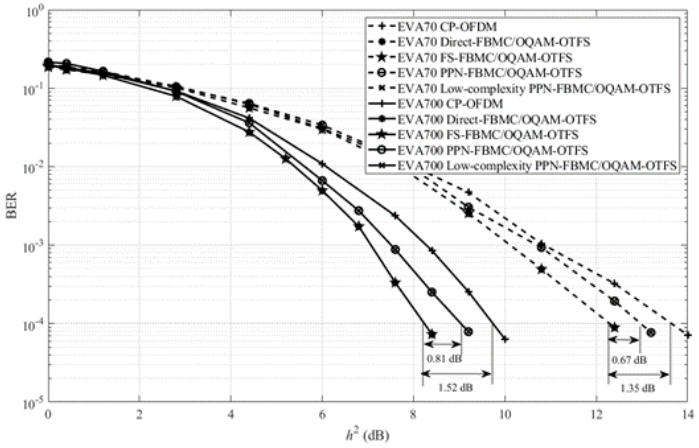


Fig. 10. BER performance over EVA 70/700 Hz of different modulation techniques

Table 5 shows the values of h^2 of mentioned modulation techniques at the BER = 10^{-4} for different multipath channel profiles EPA, EVA and ETU from moderate to highly dynamic scenarios. From Table 5 it follows that the relative positions of the BER curves in Fig. 10 remain the same, only the absolute values of h^2 change. For PPN-FBMC/OQAM-OTFS and low-complexity PPN-FBMC/OQAM-OTFS signals, the energy losses are approximately about 0.57, 0.62, 0.68, 0.8, 0.41 and 0.95 dB compared to Direct-FBMC/OQAM-OTFS and FS-FBMC/OQAM-OTFS signals in EPA5, EPA500, EVA70, EVA700, ETU300 and ETU3000, respectively. For CP-OFDM signal, the energy losses are approximately about 1.12, 1.22, 3.36, 1.51 and 3.16 compared to Direct-FBMC/OQAM-OTFS and FS-FBMC/OQAM-OTFS signals in EPA5, EPA500, EVA70, EVA700 and ETU300, respectively. CP-OFDM signal is not reached BER = 10^{-4} even on the noise-free channel, h^2 = inf.

Table 5

The energy cost of modulation techniques at the BER = 10^{-4}

	Direct-FBMC/ OQAM-OTFS	FS-FBMC/ OQAM-OTFS*	PPN-FBMC/ OQAM-OTFS*	Low-complexity PPN-FBMC/ OQAM-OTFS*	CP-OFDM*
EPA5	22.63	0	0.57	0.57	1.12
EPA500	14.25	0	0.62	0.62	1.22
EVA70	12.29	0	0.68	0.68	1.35
EVA700	8.23	0	0.8	0.8	1.52
ETU300	7.19	0	0.41	0.41	3.16
ETU3000	9.98	0	0.95	0.95	–**

* The energy loss relative to Direct-FBMC/OQAM-OTFS at BER = 10^{-4}.
** BER = 10^{-4} is not reached even on the noise-free channel, h^2 = inf.

Conclusion

This paper introduces a low-complexity FBMC/OQAM-OTFS scheme based on a polyphase structure (PPN-FBMC/OQAM-OTFS) for future wireless communication systems. Simulation results demonstrate that the proposed PPN-FBMC/OQAM-OTFS approaches achieve a reduction in computational complexity about 2.5–4 times, with an energy loss of no more than 1 dB under various channel models with different maximum Doppler shifts, when compared to the FS-FBMC/OQAM-OTFS technique. These findings suggest that the PPN-FBMC/OQAM-OTFS technique is a feasible and promising alternative to conventional OFDM in high-mobility wireless scenarios.

REFERENCES

1. **Hadani R., Rakib S., Tsatsanis M., Monk A., Goldsmith A.J., Molisch A.F.** Orthogonal time frequency space modulation. *2017 IEEE Wireless Communications and Networking Conference* (*WCNC*), 2017, Pp. 1–6. DOI: 10.1109/WCNC.2017.7925924

2. **He X., Fan P., Wang Q.** A two-stage channel estimation algorithm for OTFS in fractional doppler channels. *IEEE Communications Letters*, 2023, Vol. 27, No. 7, Pp. 1839–1843. DOI: 10.1109/LCOMM.2023.3270296

3. **Nguyen P.T.H., Khuc B., Petrov I., Lavrukhin T., Gelgor A.** Improvement in data transmission efficiency in mobile 5G new radio system using filter bank multicarrier signals. *2022 International Conference on Electrical Engineering and Photonics* (*EExPolytech*), 2022, Pp. 63–66. DOI: 10.1109/EExPolytech56308.2022.9950816

4. **Hammoodi A., Audah L., Taher M.A.** Green coexistence for 5G waveform candidates: A review. *IEEE Access*, 2019, Vol. 7, Pp. 10103–10126. DOI: 10.1109/ACCESS.2019.2891312

5. **Khuc B., Gelgor A., Le Duc T., Nguyen P.T.H.** FBMC/OQAM with OTFS pre-processing for high-mobility channels. *2024 International Conference on Electrical Engineering and Photonics* (*EExPolytech*), 2024, Pp. 186–189. DOI: 10.1109/EExPolytech62224.2024.10755876

6. **Pereira Junior R., da Rocha C.A.F., Chang B.S., Le Ruyet D.** A two-dimensional FFT precoded filter bank scheme. *IEEE Transactions on Wireless Communications*, 2023, Vol. 22, No. 11, Pp. 8366–8377. DOI: 10.1109/TWC.2023.3262442

7. **Husam A.-A., Kollár Z.** Complexity comparison of filter bank multicarrier transmitter schemes. *2018 11th International Symposium on Communication Systems, Networks & Digital Signal Processing* (*CSNDSP*), 2018, Pp. 1–4. DOI: 10.1109/CSNDSP.2018.8471795

8. **Mattera D., Tanda M., Bellanger M.** Analysis of an FBMC/OQAM scheme for asynchronous access in wireless communications. *EURASIP Journal on Advances in Signal Processing*, 2015, Vol. 23, Pp. 1–22. DOI: 10.1186/s13634-015-0191-4

9. **Bringham E.O.** *The fast Fourier transform and its applications*. New Jersey: Prentice Hall Inc., 1974.

10. **Varga L., Kollár Z.** Low complexity FBMC transceiver for FPGA implementation. *2013 23rd International Conference Radioelektronika* (*RADIOELEKTRONIKA*), 2013, Pp. 219–223. DOI: 10.1109/RadioElek.2013.6530920

11. **Vo-Huu T., Ludant N., Vo-Huu T., Noubir G.** Pilotless FS-FBMC for flexible spectrum access and sharing. *2021 IEEE International Symposium on Dynamic Spectrum Access Networks* (*DySPAN*), 2021, Pp. 193–202. DOI: 10.1109/DySPAN53946.2021.9677265

INFORMATION ABOUT AUTHORS / СВЕДЕНИЯ ОБ АВТОРАХ

Bang Thanh Khuc
Хук Тхань Банг
E-mail: khucbang@mail.ru

Aleksandr L. Gelgor
Гельгор Александр Леонидович
E-mail: agelgor@spbstu.ru

Submitted: 22.09.2025; Approved: 20.11.2025; Accepted: 13.12.2025.
Поступила: 22.09.2025; Одобрена: 20.11.2025; Принята: 13.12.2025.

Computing, Telecommunication and Control, 2025, Vol. 18, No. 4, Pp. 67–75.
Информатика, телекоммуникации и управление. 2025. Том 18, № 4. С. 67–75.

Research article
DOI: https://doi.org/10.18721/JCSTCS.18406
UDC 621.397

DESIGN OF IOT DEVICE USING BEAM-SPLITTING PRISM DISPLAY

Xu Luolan ✉

Xi'an Jiaotong University, Xi'an, China

✉ xuluolan123@gmail.com

Abstract. This project is based on an open-source hardware. Its functions have been redesigned and improved, with enhancements made to the hardware programming and power supply circuits. System-level secondary development has been carried out to implement an embedded system interface and a weather clock application. During development, modifying device functionality required close coordination between software and hardware and optimization of the enclosure structure. As resource demands increases, adjustments to the hardware power supply and software optimization become necessary to ensure reliable operation. Internet of things (IoT) device development necessitates a holistic approach. By working backward from the target specifications, hardware, software and enclosure design must be considered together. The results indicate that optimized power supply, low-coupling software operation and a thermally efficient enclosure significantly enhance the long-term stability and low-power operation of IoT devices.

Keywords: embedded systems development, internet of things (IoT), devices, beam-splitting prism display, low-power optimization, graphical user interface (GUI)

Acknowledgements: I would like to thank Peng Zhihui, the author of the open-source project, the users of the project's forum, my university and its teaching staff for their assistance in the implementation of this research.

Citation: Xu Luolan. Design of IoT device using beam-splitting prism display. Computing, Telecommunications and Control, 2025, Vol. 18, No. 4, Pp. 67–75. DOI: 10.18721/JCSTCS.18406

Научная статья
DOI: https://doi.org/10.18721/JCSTCS.18406
УДК 621.397

РАЗРАБОТКА УСТРОЙСТВА IOT С ИСПОЛЬЗОВАНИЕМ ДИСПЛЕЯ НА ОСНОВЕ ЛУЧЕПРЕЛОМЛЯЮЩЕЙ ПРИЗМЫ

Л. Сюй ✉

Сианьский транспортный университет, Сиань, Китай

✉ xuluolan123@gmail.com

Аннотация. Настоящий проект основан на аппаратном обеспечении с открытым исходным кодом. Его функции были переработаны и улучшены, включая усовершенствование схем аппаратного программирования и питания. Была проведена дополнительная разработка на системном уровне для реализации встроенного системного интерфейса и программы для метеостанции с часами. В ходе разработки модификация функциональности устройства потребовала тесной координации между программным и аппаратным обеспечением, а также оптимизации конструкции корпуса. При увеличении потребления ресурсов для обеспечения надежной работы потребовалась настройка аппаратного обеспечения и оптимизация программного обеспечения. Разработка устройств Интернета вещей (IoT) требует комплексного подхода. Исходя из целевых технических характеристик, необходимо рассматривать в комплексе аппаратное, программное обеспечение и дизайн корпуса. Результаты показывают, что оптимизированный источник питания, программное обеспечение с низким энергопотреблением и термоэффективность корпуса значительно увеличивают срок службы и улучшают стабильность работы IoT-устройств.

Ключевые слова: разработка встраиваемых систем, устройства, интернет вещей (IoT), дисплей на основе лучепреломляемой призмы, оптимизация энергопотребления, графический интерфейс пользователя (GUI)

Благодарности: Я хотел бы поблагодарить Пэн Чжихуэя, автора проекта с открытым исходным кодом, пользователей форума проекта, мой университет и его преподавательский состав за помощь в реализации данного исследования.

Для цитирования: Xu Luolan. Design of IoT device using beam-splitting prism display // Computing, Telecommunications and Control. 2025. Т. 18, № 4. С. 67–75. DOI: 10.18721/JCSTCS.18406

Introduction

With the rapid development of information technology, the proliferation of Internet of things (IoT) devices has permeated various aspects of personal life, such as smart home and smart in-car systems. Therefore, the rapid development of versatile IoT interactive terminal devices capable of connecting to various gateways holds significant research value and importance [1].

This paper explores the secondary development of an open-source beam-splitting prism display device, enabling beginners to learn and master hardware design and optimization, as well as secondary software development and comprehensive understanding of embedded systems [2].

Progress in electronics often comes from standing on the shoulders of giants. Open-source projects represent the cumulative effort of numerous developers gradually refining a design. This study focuses on creating an IoT device based on ESP32 hardware circuits and beam-splitting prism display, involving circuit optimization, program modification, and functional enhancements such as temperature control and weather information display. By creating an actual application, the device can later be extended into a multifunctional integrated system – hence the project name AIO (All-In-One).

Research design

The ESP32-PICO-D4 was selected as the core controller for this research. It features a dual-core 32-bit processing core, RTC and low-power management module [3], and provides complete Wi-Fi and Bluetooth functions, which greatly enhances the device's practicality.

For the display, given the current trend toward transparent frameless screens, a balance between technology and cost-effectiveness was sought. The method of prism refraction display can be used to present various functions in a "transparent" manner at low cost. This project's name AIO emphasizes that the hardware platform serves as a foundation upon which a wide range of functions – such as the weather and time display in this implementation – can be added through secondary software development.

Hardware design

Based on the open-source hardware, this project focuses on implementing weather and time display with network NTP synchronization. The original hardware design supported only basic image display functions, resulting in relatively low power consumption. However, the long-term simultaneous operation of network time synchronization, screen display, accelerometer, RGB lights and other components will greatly increase the overall circuit power consumption, requiring modifications to the hardware power supply [4].

The updated hardware design is as follows.

A two-layer printed circuit board (PCB) design was adopted. The upper PCB is designed for screen display and optimized power supply. According to the size of the prism, a 1.3-inch IPS color display with a 240×240 resolution (square format) was selected, ensuring the display area aligns precisely with the prism's refracting region. The lower PCB integrates main controller, accelerometer, RGB lights and TransFlash card slot. The two circuit boards are connected by flexible flat cable (FFC) connectors, which provide power and signal communication. Fig. 1 shows the front view of the lower main control board of the device.

For the upper voltage-stabilizing display module of the device, the original circuit's voltage-stabilizing chip failed to provide stable power supply, and it was replaced with the ME6211 linear voltage regulator to supply power to the entire device. Fig. 2 shows the front-side layout of the voltage-regulator and display board.

Since there are no physical buttons, all control and interaction rely on the MPU6050 accelerometer, which is placed in the middle of the lower main control board circuit to ensure accurate parameters during program initialization. The ESP32 and the CP2102 chip (used for programming)

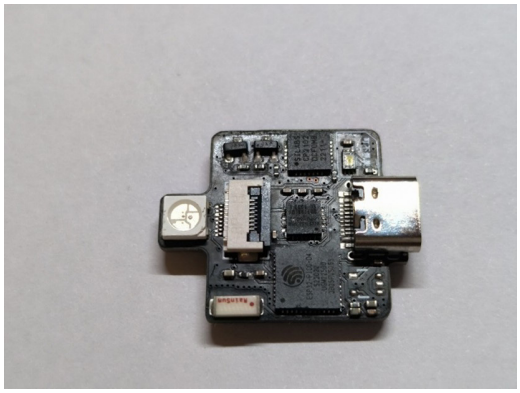


Fig. 1. Lower main control board

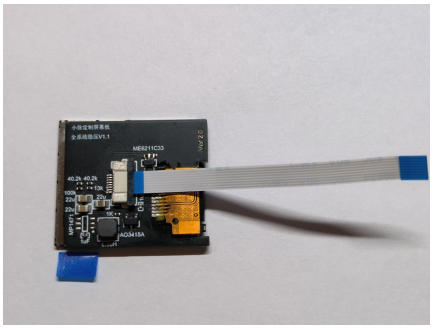


Fig. 2. Front-side layout of the voltage-regulator and display board

are placed on the upper and lower sides of the accelerometer, respectively, while the Type-C port and RGB LEDs are placed on the left and right sides. The back of the main control board is designed with a TF card slot for storing images and videos in later stage, as well as files for development and calling. This helps to save the memory space of the MCU itself.

The original project used the LP2992 linear voltage regulator, which supports a maximum output current of 250 mA. However, the screen alone requires approximately 100 mA when illuminated. Moreover, subsequent maintenance programs occupy a large amount of processing capacity, leading to excessive power consumption that can cause severe heating of the low-dropout (LDO) regulator and even program crashes. To solve this problem, this project replaced the ME6211C33 linear voltage regulator chip in the screen's power supply board, increasing the load current capacity to 500 mA and thereby resolving the issue of overheating of the original IC. The optimized parameters of the updated circuit are shown in Table 1.

Table 1

Comparison of updated device parameters

	Original	**Modified**
Performance parameters	LP2992 3V3	SGM662K-3.3
Maximum output current	250 mA	**500 mA**
Typical dropout voltage	115 mV @ 100mA 200 mV @ 150mA	**130 mV @ 300 mA** **200 mV @ 500 mA**
Quiescent current	Typical 170 μA	**Typical 30 μA**
Ripple rejection ratio	75 dB @ 1kHz 45 dB @ 100kHz	75 dB @ 1 kHz 40 dB @ 10 kHz
Protection functions	**Overcurrent, overtemperature, reverse current protection, enable control**	Overcurrent, overtemperature, reverse current protection
Input voltage range	**2.5V ~ 16V**	2.2V ~ 5.5V
Operating junction temperature range	**–40°C ~ +125°C**	–40°C ~ +85°C
Circuit ripple	22 mV	**18 mV**
Cost	Relatively high	Very low

It can be seen that the updated power supply IC provides optimization in terms of low power consumption, high-current drive capability and cost-effectiveness. Additionally, actual tests show that the power ripple across the overall power supply has been reduced.

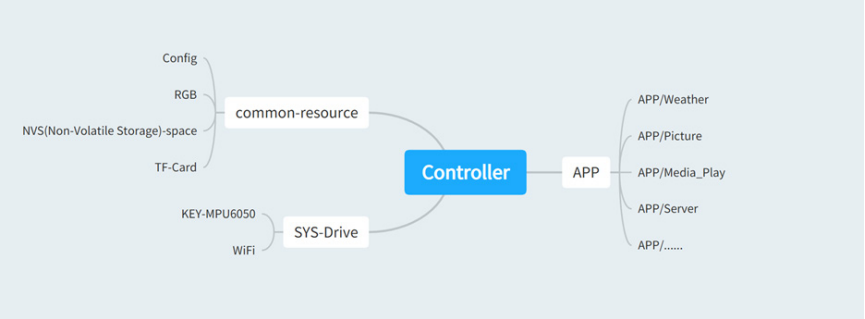


Fig. 3. System design mind map

Software framework

Based on the hardware design, the software must account for the actual computing power of the MCU. Therefore, the light and versatile graphics library (LVGL) was selected to handle the control display interface, which also helped shorten the software development cycle. The logical design of the system framework can be understood from the overall system design mind map shown in Fig. 3.

It can be seen that all functions are implemented through the main control class (system graphical interface). Since the hardware uses an accelerometer for control, the sensor's output is processed as the key value to realize human-machine interaction logic, which is the core control method of the device – operation via the gyroscope sensor. At the software level, the key value can be used as a controller to select the running app code module, thereby realizing different functions [5]. Once inside the app, the key value is also used for control or selection. In this way, the integrated sensor fully replaces various external devices such as computer mouse or keyboard, making the device very compact and fully functional – consistent with the design philosophy of small, portable IoT devices.

Each app (structured as a module) can access public resources as long as it meets the interface specifications of the overall framework. For example, access to the TF card, RGB lights, global configuration files, network settings and small system parameters can all be modified and controlled. This low-coupling program design enables rapid customization of software according to existing hardware resources during app development.

It should be noted that, due to the physical characteristics of the light-splitting prism, during the system development process, the image displayed on the screen must be mirrored to display the correct content that appears on the front side of the prism.

Implementation of weather system app function

On the 1.3-inch color display, in addition to text data, weather icons can be added to enhance the visual effect [6]. This requires the use of the display font library and the integration of the LVGL [7]. Since this project only uses weather and region-related content, other elements of the library can be removed to save processor storage space.

For example, Chinese Amap Open Platform provides real-time updates of temperature, humidity, wind strength and air quality for specified cities, as well as weather forecast information for the next week. The app ID, app secret , city name, and network synchronization interval information are recorded in the ESP32 to avoid repeated settings after power loss or restart [8].

However, since weather conditions seldom change over short periods of time, displaying this information is unnecessary. Moreover, this may lead to screen burn-in. Therefore, this project incorporates dynamic clock display. On one hand, it can update data in real time to prevent screen burn-in; on the other hand, it enhances the visual appeal of the weather app.

The weather app consists of the following modules, coordinated via functions or messaging mechanisms:

1. Initialization Phase
   a. Read or write weather-related configurations (e.g., city name, update interval).
   b. Prepare basic styles such as fonts and color schemes for the interface.
   c. Initialize runtime data for storing weather and time information.
2. Main Loop/Event Processing
   a. Continuously monitor user input (e.g., return, left/right switching) and system messages (e.g., Wi-Fi connection, parameter settings) in a loop or timed callbacks.
   b. Determine the next action based on the current page type (weather or curve page): update weather data or time, or simply display existing data.
3. Data Acquisition and Parsing
   a. When the network is available, access the corresponding weather and time API; parse JSON responses or timestamps.
   b. When the network is unavailable, use local millisecond counters for time calculation until connection is restored.
   c. Store the obtained or calculated results in the runtime data structure.
4. Interface Display and Refresh
   a. Update graphical interface when new data is available.
   b. On the weather page: display current temperature, humidity, wind strength, city name, etc., along with animations.
   c. On the curve page: draw weekly maximum/minimum temperature curve.
   d. When the network is available, update the clock display (hours, minutes, seconds, date).
5. Exit/Cleanup
   a. When the user or system requests to exit, destroy interface objects and styles, close background tasks, release memory and ensure that no system resources are retained.

The flowchart of the weather program design is shown in Fig. 4. It should be noted that the graphical interface of the weather app is developed based on LVGL, including interface and style, multi-page management, weather display, forecast curve, clock and date display, animation effects, etc. This project draws inspiration from the clock interface template made by Misaka. Image design is customizable according to personal preferences and will not be described in detail in this paper.

Shell design

Since the designed IoT device needs to run continuously and the ESP32 periodically performs network time calibration through Wi-Fi, heat generated by the power supply and screen display will inevitably accumulate [3]. If heat dissipation cannot be guaranteed, the operation stability of the IoT device will be reduced [9]. Therefore, the shell connecting the main control PCB, screen PCB and upper light-splitting prism will assume the main heat dissipation function. In the early stage, 3D printing can be used to determine the rationality of the hardware structure distribution. In the later stage, the designed structure can be made of aluminum alloy, and the shell model parameters can be delivered to a factory capable of CNC machining [10]. Thanks to technological advancements and the convenience of online shopping, the cost of small-batch design and processing has been greatly reduced. It should be noted that due to the adoption of a metal heat dissipation shell design, the PCB should use the via tenting process, and during assembly, the metal parts on the circuit board must be properly insulated for the circuit.

Results

After powering on the assembled device, it first boots into the system's main interface. From there, the weather app can be selected and entered through tilt-based navigation. After entering the interface,

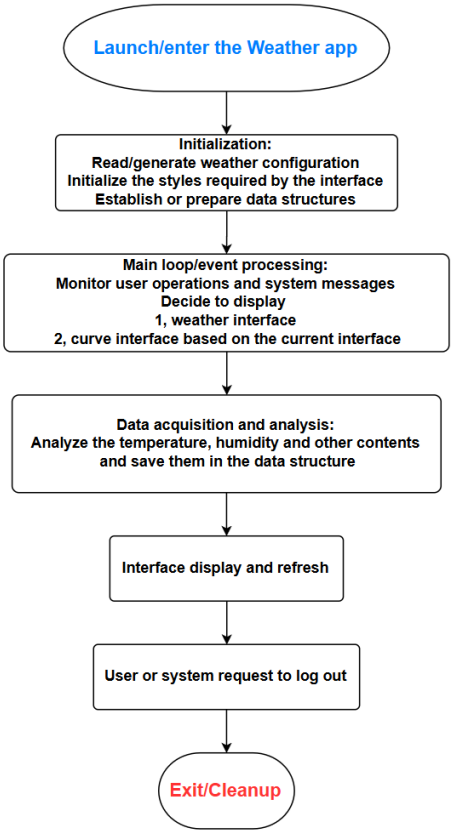


Fig. 4. Weather app design flowchart

the program initially displays the default time and weather screen, and the time is updated according to the RTC built into the ESP32, which ensures normal app operation. Shortly after, the device experiences a brief lag, while the program is calibrating the weather and time data over the network. Subsequently, the interface will display the synchronized information, indicating that the overall design functions are operational. After 24 hours of continuous operation, as shown in Fig. 5, the displayed time remains accurate to the second due to regular calibration, with no observed desynchronization. This also demonstrates stable performance of the device's network and weather display functions.

By touching the metal shell part, it is evident that most of the heat generated by the device is transferred to the shell, and the temperature is stably controlled at a level slightly above ambient level, indicating that the effective auxiliary heat dissipation is provided by the shell.

Finally, all content is presented in the transparent glass through the refraction and reflection of the prism.

Conclusion

Through the design and implementation of this prism-based display IoT project, the intended functions have been realized via integrated hardware and software design, shell design, and novel prism display design. A deeper understanding of the comprehensive optimization of device stability has been obtained. This project-based learning approach has proven to be a highly effective methodology.

During hardware design, it was found that after software updates, the screen failed to display properly due to insufficient current; the power supply was subsequently modified. During software design, app development was carried out using a low-coupling approach to reduce redundant code. During shell design, overheating was found to cause system crashes, and the shell material was improved to

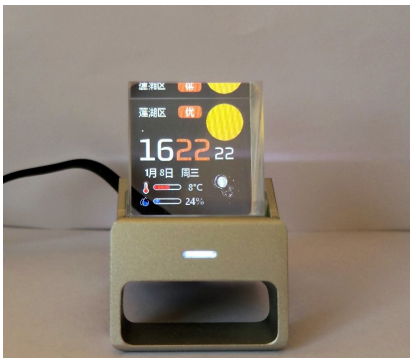


Fig. 5. Equipment operation

enhance overall thermal conductivity. These modifications were identified and implemented during the secondary development of the original device.

Compared with the initial hardware design, the combination of optimized voltage-stabilizing circuit, metal shell and main control board heat dissipation realizes more stable voltage control, thereby improving the stability of the module during long-term use and reducing circuit ripple noise. At the software level, by setting public reusable interfaces, subsequent development on this hardware platform can incorporate customized apps within the overall framework. The IoT module can adapt to different usage scenarios by software upgrades, significantly reducing the app development cycle of the underlying operating system.

In the future, the number of IoT devices is expected to show explosive growth and will appear in all aspects of people's lives in various forms [8]. In terms of interaction logic, design will continue to engage human visual, motion, auditory, and tactile perception [5]. Presenting the interface in a more elegant manner remains a persistent goal for engineers. Based on the design of the minimum core module of the IoT, this project presents an interactive form different from traditional screen display. It is hoped that through this way, open-source IoT devices can show a more diversified development.

REFERENCES

1. **Jocknoi L., Kucharoen P.** ESP32Exten: Designing and developing an ESP32 microcontroller expansion for IoT applications with motor propulsion and AI image processing. *2024 8th International Conference on Information Technology* (*InCIT*), 2024, Pp. 278–283. DOI: 10.1109/InCIT63192.2024.10810578

2. **Wang Z., Tu K., Lv G., Feng Q.** Depth enhanced holographic super multi-view display based on multiple image recording planes. *IEEE Journal of Selected Topics in Quantum Electronics*, 2024, Vol. 30, No. 2, Art no. 6000506. DOI: 10.1109/JSTQE.2024.3364581

3. **Farid N.A.M., Razak A.H.A., Halim A.K., Idros M.F.M.** Design of CMOS RF Doherty power amplifier in low-power 5G wireless networks for IoT application. *2023 IEEE 11th Conference on Systems, Process & Control* (*ICSPC*), 2023, Pp. 310–314. DOI: 10.1109/ICSPC59664.2023.10420155

4. **Saxena A., Haripriya D., Madan P., Srivastava A.P., Shalini N., Kumar A.** Design and optimization of low-power VLSI circuits for IoT devices. *2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering* (*UPCON*), 2023, Pp. 1267–1273. DOI: 10.1109/UPCON59197.2023.10434775

5. **Nwadiugwu W.P., Kim D.-S.** Energy-efficient sensors in data centers for Industrial Internet of Things (IIoT). *2018 3rd International Conference on Internet of Things: Smart Innovation and Usages* (*IoT-SIU*), 2018, Pp. 1–6. DOI: 10.1109/IoT-SIU.2018.8519871

6. **Devi A., Arivunambi A., Suvetha S., Sasikala S., Dharanyadevi P., Senthilnayaki B.** IoT based satellite balloon system for live-weather forecast. *2024 3rd International Conference on Smart Technologies and Systems for Next Generation Computing* (*ICSTSN*), 2024, Pp. 1–5. DOI: 10.1109/ICSTSN61422.2024.10671106

7. **Zaharia S., Rebedea T., Trausan-Matu S.** Source code vulnerabilities detection using loosely coupled data and control flows. *2019 21st International Symposium on Symbolic and Numeric Algorithms for Scientific Computing* (*SYNASC*), 2019, Pp. 43–46. DOI: 10.1109/SYNASC49474.2019.00016

8. **Cheng Z.** Research on Internet of Things human-computer interaction system based on computer Artificial Intelligence technology. *2024 IEEE 2nd International Conference on Control, Electronics and Computer Technology* (*ICCECT*), 2024, Pp. 1135–1139. DOI: 10.1109/ICCECT60629.2024.10545728

9. **Wijittemee W., Plangklang B.** An efficient thermal management using passive cooling techniques. *2024 International Conference on Power, Energy and Innovations* (*ICPEI*), 2024, Pp. 73–77. DOI: 10.1109/ICPEI61831.2024.10748605

10. **Boora A., Verma D., Bijender, Soni A.** Time reduction analysis based on infill pattern on FDM 3D printed PLA material. *2024 International Conference on Intelligent & Innovative Practices in Engineering & Management* (*IIPEM*), 2024, Pp. 1–4. DOI: 10.1109/IIPEM62726.2024.10925750

INFORMATION ABOUT AUTHOR / СВЕДЕНИЯ ОБ АВТОРЕ

Xu Luolan
Сюй Лолань
E-mail: xuluolan123@gmail.com

Submitted: 10.01.2025; Approved: 13.11.2025; Accepted: 09.12.2025.
Поступила: 10.01.2025; Одобрена: 13.11.2025; Принята: 09.12.2025.

Computing, Telecommunication and Control, 2025, Vol. 18, No. 4, Pp. 76–86.
Информатика, телекоммуникации и управление. 2025. Том 18, № 4. С. 76–86.

Software and Hardware of Computer, Network, Telecommunication, Control, and Measurement Systems

Компьютерные сети, вычислительные, телекоммуникационные, управляющие и измерительные системы

Research article
DOI: https://doi.org/10.18721/JCSTCS.18407
UDC 004.312.44

DESIGN AND ANALYSIS OF A RECONFIGURABLE HARDWARE ACCELERATOR FOR SOLVING A SYSTEM OF LINEAR EQUATIONS USING JACOBI METHOD

M.F. Gonzalez, A.P. Antonov ✉

Peter the Great St. Petersburg Polytechnic University,
St. Petersburg, Russian Federation

✉ antonov_ap@spbstu.ru

Abstract. This work presents the design and analysis of a reconfigurable hardware accelerator for solving a system of linear equations using Jacobi method, implemented on a reconfigurable device, as well as a comparative study of software and hardware implementations. Recent advancements in computing capabilities have been hindered by the so-called "walls": memory, power consumption and clock frequency limitations imposed by current technology. Solutions to overcome these "walls" include reconfigurable computing and high-level synthesis. The system under development and analysis was described in the C++ language and implemented using a high-level synthesis method, which reduces design time and enables more efficient exploration of different hardware architectures. The comparative analysis showed a performance increment over the original implementation, with energy consumption comparable to that of a modern mid-class microprocessor.

Keywords: supercomputer, reconfigurable hardware accelerator, Jacobi method, field-programmable gate array, high-level synthesis, SystemVerilogHDL, reconfigurable computing

Citation: Gonzalez M.F., Antonov A.P. Design and analysis of a reconfigurable hardware accelerator for solving a system of linear equations using Jacobi method. Computing, Telecommunications and Control, 2025, Vol. 18, No. 4, Pp. 76–86. DOI: 10.18721/JCSTCS.18407

Научная статья
DOI: https://doi.org/10.18721/JCSTCS.18407
УДК 004.312.44

ПРОЕКТИРОВАНИЕ И АНАЛИЗ РЕКОНФИГУРИРУЕМОГО АППАРАТНОГО УСКОРИТЕЛЯ ДЛЯ РЕШЕНИЯ СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ МЕТОДОМ ЯКОБИ

М.Ф. Гонзалез, А.П. Антонов ✉

Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Российская Федерация

✉ antonov_ap@spbstu.ru

Аннотация. В данной работе представлены разработка и анализ реконфигурируемого аппаратного ускорителя для решения системы линейных уравнений методом Якоби, реализованного на реконфигурируемом устройстве. Проведено сравнительное исследование производительности аппаратной и программной реализаций. В последнее время рост вычислительных возможностей высокопроизводительных вычислительных систем сдерживается такими преградами, как память, энергопотребление, тактовая частота, накладываемыми современными технологиями. Решениями для преодоления указанных преград являются реконфигурируемые вычисления и высокоуровневый синтез. Разрабатываемая и анализируемая система была описана на языке C++ и реализована с использованием метода высокоуровневого синтеза, что позволило сократить время проектирования и эффективнее исследовать различные аппаратные архитектуры. Сравнительный анализ показал увеличение производительности по сравнению с первоначальной реализацией при потреблении энергии, сопоставимом с современным микропроцессором среднего класса.

Ключевые слова: суперкомпьютерный вычислитель, реконфигурируемый аппаратный ускоритель, метод Якоби, сверхбольшие интегральные схемы программируемой логики, высокоуровневый синтез, язык SystemVerilogHDL, реконфигурируемые вычисления

Для цитирования: Gonzalez M.F., Antonov A.P. Design and analysis of a reconfigurable hardware accelerator for solving a system of linear equations using Jacobi method // Computing, Telecommunications and Control. 2025. Т. 18, № 4. С. 76–86. DOI: 10.18721/JCSTCS.18407

Introduction

The computing capabilities of microprocessors increased steadily for decades. However, this trend has recently been significantly disrupted due to the so-called “walls”: memory, power consumption, clock frequency/technology. As an alternative to general-purpose microprocessors for performance-critical tasks, application-specific integrated circuits (ASICs) can be used. However, they involve high complexity, cost and design time on the one hand, and narrow specialization for the solution of a single task or the implementation of one algorithm, on the other hand. The solution of general problems using graphics processing units (GPGPUs), which fall into the single instruction, multiple data (SIMD) category according to Flynn’s taxonomy, is not efficient in terms of performance and power consumption, since GPGPU is optimized for vector and vectorizable problems [1–3].

This trend is evident, for example, in the analysis of the Top 500 list. On the one hand, performance growth can be observed, achieved by increasing the number of cores and power consumption. On the other hand, as showed in [4], using the conjugate gradient method instead of the “classical” Linpack leads to performance degradation of two orders of magnitude.

Thus, there is a relationship between performance and computing architecture: existing architectures, optimized for solving one task, are not efficient for solving similar tasks with different methods, let alone tasks of different classes.

A solution to this problem is the use of hardware reconfigurable accelerators, whose physical structure is adapted to the algorithm of the task to be solved [5, 6].

The base of reconfigurable accelerators is hardware reconfigurable devices, such as field-programmable gate arrays (FPGAs). Widely known FPGAs enable the implementation of hardware solutions for computationally complex tasks with high performance, close to that of ASICs, but with significantly lower time and resource consumption. However, the traditional approach to FPGA design, including the use of hardware description languages, require specialized knowledge and skills, which has essentially limited the practical adoption of FPGAs as reconfigurable accelerators.

A solution to this problem is the use of high-level synthesis methods [7], which enable an abstraction of the complexity of hardware development. Apart from the automatization of the design process, high-level development environments enable the generation and analysis of a wide range of hardware architectures, including parallel and pipelined variants.

The system of linear equations (SLE) models a broad spectrum of scientific problems, such as weather forecasting or finite element methods. SLE is typically represented in the form of matrix-vector multiplication, which is efficiently processed by computers. There are different approaches to solving SLE, mainly direct and iterative methods. This article focuses on the Jacobi iterative method.

Unlike other methods, such as Gauss–Seidel, the Jacobi method computes all equations before updating the values of the unknown, meaning that hard dependencies between variables arise only between computations at different iterations. This characteristic gives the Jacobi method a high degree of parallelization, which increases with the size of the problem [8]. For this reason, a massively parallel architecture may represent an effective platform for solving SLE using the Jacobi iterative method. Thus, reconfigurable devices appear particularly well-suited for this task.

Several authors have presented their contributions to this problem in recent years [9–11]. A general formulation of the problem is presented at [12]. As in most hardware implementations, there are several possible approaches, from completely parallel systems to fully sequential ones, as well as intermediate levels of parallelism. In [12], both parallel and sequential approaches are proposed, while also outlining the principles of pipelined processing systems. Considerations regarding system complexity are also discussed.

In [13], the Jacobi method is applied to fractal calculations. According to the finite element method, each element in a mesh is most closely related to the neighboring nodes. These dependencies are modeled by SLE, where each equation computes a new value for each node according to the values of its four nearest neighbors. A convenient degree of parallelism is achieved by dividing the mesh into sub-frames, which are computed in parallel.

In [14], a similar problem is addressed, related to solving a large SLE using the Jacobi iterative method. In this case, the mathematical problem under investigation is solving Laplace's equations with Dirichlet boundary conditions, where standard (non-boundary) nodes are computed based on their four neighbors. This structure is modeled by a computing system based on computing blocks (nodes) interconnected by communication channels. The connections may be physical, corresponding to channels of spatially parallel structures, or virtual, represented by the exchange of information between iterations.

The computing architecture is also studied from different points of view: as fully parallel (spatial), fully sequential (temporal) or as a combination of both approaches adapted to the computing resources available on the reconfigurable device. The elementary computing block may receive input data in parallel, which in general corresponds to the values of neighboring nodes, or simpler but slower computing blocks may receive this data sequentially, processing it using corresponding accumulators and registers for storing the intermediate calculations between clock cycles. This reduction in the required logical resources simultaneously allows to limit the number of communication channels to a small fraction of the initially needed amount.

The decision on which approach will be implemented depends on the ratio of required computing resources, related to the size of the problem, to the resources available on the physical device. Therefore, elementary blocks, whether based on accumulators or not, may be reproduced several times according to the ratio of required to available resources, taking into account the achieved performance. A computing structure pipelined in several stages allows achieving higher performance without a significant increase of required logical resources by their reuse once intermediate results are no longer needed. The limited number of available communication channels also plays a key role, since the ratio of external communication channels on reconfigurable devices to the increasing area of logical computing elements has been observed to decrease constantly. The authors also studied the performance advantage of using clusters of reconfigurable devices for solving large-scale problems.

It is commonly accepted that the traditional approach based on hardware description language (HDL) design is excessively complex and time-consuming, especially for large systems of equations that required to be adapted to the resources available on the target device. In [11], a design for a Maxeler acceleration card is presented, using MaxCompiler, which allows system to be described in Java language. The key proposal for pipelining the system is to compute as many equations as necessary within the first iteration so that the second and successive iterations may start afterwards in a pipelined fashion, overcoming the inter-iteration dependencies described above.

In [15], the Jacobi iterative method in a high-level language for hybrid reconfigurable platforms is described. In [16], the Jacobi solver on a SCR heterogeneous supercomputer using the HLL-to-HDL compiler is implemented. Mapping the floating-point units was identified as the main difficulty faced by the authors. In [17], the use of Jacobi solvers is proposed, each solving one of the equations and acting in parallel within the global system. Other related architectural aspects, aimed at increasing the performance, are presented in [18], taking into account the use of accumulators for computing the exit condition, as soon as a sufficiently precise result is achieved.

Materials and methods

The object of the research is a method for increasing the performance of solving SLE.

SLE is typically presented in the following way:

$$\begin{gathered} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = b_1; \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n = b_2; \\ \cdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,n}x_n = b_n. \end{gathered} \tag{1}$$

An algebraic representation of SLE follows:

$$Ax = p. \tag{2}$$

As mentioned above, SLE may be solved using either direct or iterative methods. Iterative methods are particularly advantageous for computational solutions. A classical iterative method is the Jacobi algorithm, based on the provision of an initial guessed approximate solution x^0, typically expected to be close to the exact solution, and on the computation of the value of each unknown assuming that the remaining unknowns in the current equation have the value from the initial vector x^0. For each equation, the unknowns placed at the diagonal of the matrix A (2) are computed using the following procedure:

$$x_i^{k+1} = \frac{\left(b_i - \sum_{i \neq j}^{n} a_{i,j} x_j^k\right)}{a_{i,j}}. \quad (3)$$

The obtained result is the approximation x^1 to the exact solution. A more accurate solution may be obtained by another computation of the values of each unknown, this time based on the previously computed vector x^1 for the remaining unknowns instead of vector x^0. The process may be repeated iteratively until a solution with an acceptably small error is found. The criteria of an admissible error are usually expressed in relation to the norm of the right-hand vector b. The method converges if, for each row (or column) of matrix A, the absolute value of the diagonal element is greater than that for each individual element in the same row (or column).

Other iterative methods include the Gauss–Seidel method, which takes each newly computed value of the unknowns and immediately uses it for computing the following equation within the same iteration. This typically leads to a faster SLE solving than the Jacobi method. However, from the computational point of view, this approach introduces data dependencies between consecutive equations within each iteration, instead of between consecutive iterations, as in the Jacobi method. Such tight dependencies significantly reduce the degree of parallelism in the algorithm. For this reason, the Jacobi method is often preferred for parallel implementations.

The subject of the research is the Jacobi algorithm for solving SLE. This algorithm was chosen due to its wide use in many applications and its high potential for parallelism.

In this work, the specific case of SLE is examined with the matrix A containing five non-zero diagonals.

$$A = \begin{pmatrix} c_1 & d_1 & \cdots & e_1 & & & & \\ b_2 & c_2 & d_2 & \cdots & e_2 & & & \\ & b_3 & c_3 & d_3 & \cdots & e_3 & & \\ a_4 & \cdots & b_4 & c_4 & d_4 & \cdots & e_4 & \\ & a_5 & \cdots & b_5 & c_5 & d_5 & \cdots & e_5 \\ & & a_6 & \cdots & b_6 & c_6 & d_6 & \cdots \\ & & & a_7 & \cdots & b_7 & c_7 & d_7 \\ & & & & a_8 & \cdots & b_8 & c_8 \end{pmatrix}. \quad (4)$$

The diagonals are arranged symmetrically with outer diagonals located at a distance L from three central diagonals. Thus, SLE may be expressed in the following way:

$$\begin{gathered} c_i x_i + d_i x_{i+1} + e_i x_{i+L} = p_i, \quad i = 1; \\ b_i x_{i-1} + c_i x_i + d_i x_{i+1} + e_i x_{i+L} = p_i, \quad i = 2, \ldots, L; \\ a_i x_{i-L} + b_i x_{i-1} + c_i x_i + d_i x_{i+1} + e_i x_{i+L} = p_i, \quad i = L+1, \ldots, N-L; \\ a_i x_{i-L} + b_i x_{i-1} + c_i x_i + d_i x_{i+1} = p_i, \quad i = N-L+1, \ldots, N-1; \\ a_i x_{i-L} + b_i x_{i-1} + c_i x_i = p_i, \quad i = N. \end{gathered} \quad (5)$$

The approximate solution is found by assigning the initial guessed value of the unknowns and the repeated iteration until the margin of error is acceptable.

Each of five diagonals may be effectively represented computationally as vectors, considered as constant data within the algorithm. The assumed values of the unknowns are variable data, updated gradually through the instructions inside the main inner loop, but are not updated until the end of each iteration.

This provides the possibility of massive parallelism for large systems of equations, due to the absence of read-after-write (RAW) dependencies within the inner loop. Nevertheless, consecutive iterations do present unsolvable RAW dependencies.

The algorithm stops when the solution error is sufficiently small. The following expression presents the condition of acceptable error:

$$\frac{\left\|b - A_x^k\right\|}{\left\|b\right\|} < \varepsilon. \tag{6}$$

The algorithm consists of two basic cyclical steps: the update of the unknowns X and the computation of the error E. Fig. 1 presents the behavior of the algorithm.

This work carries out the analysis of existing implementations, aimed at achieving the best performance, leading to the conclusion that a hardware implementation of the Jacobi iterative method using reconfigurable accelerators is highly relevant.

This work employs the following *research methods*:

- simulation modeling of software and hardware implementations for solving SLE using the Jacobi method;
- comparative analysis of hardware performance and hardware costs.

The goal of the analysis and synthesis of hardware implementation of the Jacobi iterative method is to achieve higher performance mainly through the parallelization of the algorithm.

Fig. 2 shows the parallelism applied to the basic conceptual computing units. For high-level language descriptions, each of the base blocks may be parallelized using the declaration of the type #pragma unroll applied to its loops.

The design may be additionally parallelized using the pipelined processing, i.e., dividing each computationally intensive instruction within the loops into several stages. In this way, intermediate results from each stage are stored in certain registers at each clock cycle, so that the next instruction may immediately start the completion of the stage accomplished by the previous instruction. This concept, similar to the pipelined processing in standard microprocessors, is shown in Fig. 3.

Considering the approximative character of the algorithm, it is possible to define asymmetric workloads, where E is not executed for each newly computed value of X, but only after a given k number of X computations. Consequently, the implementation is more effective when hardware resources are allocated more heavily to computing X than to computing E.

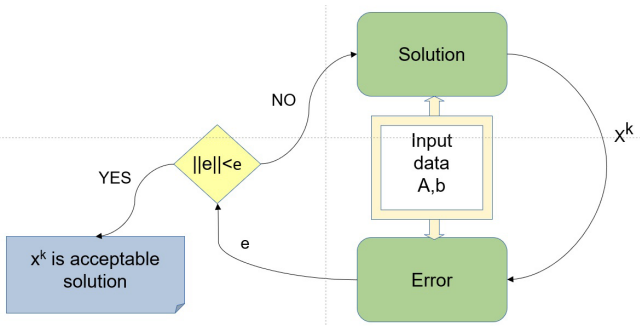


Fig. 1. Behavior of the described algorithm

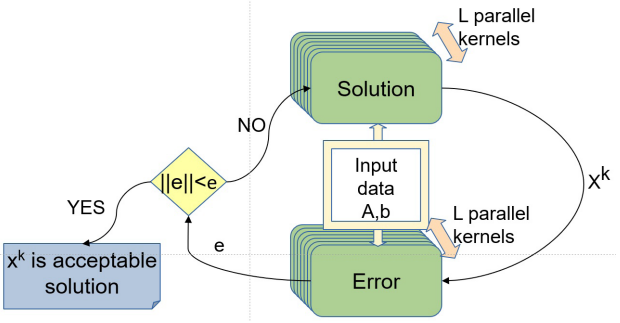


Fig. 2. Parallel approach representation

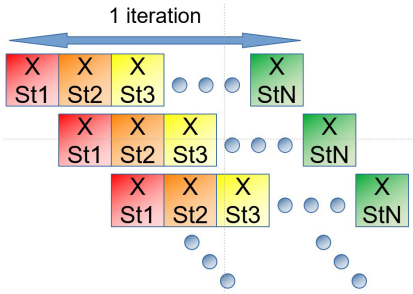


Fig. 3. Pipeline representation

An analysis of several approaches for the implementation of the elementary loops of the algorithm was carried out, in particular, to computing X. Assuming a direct translation of the algorithm to hardware resources, each equation may be represented by a line of code inside the inner loop of the algorithm, each line synthesized into hardware resources performing the required multiplications, subtractions, additions and divisions. Considering that for large problems the size N of the Jacobi matrix is usually much larger than the distance L of the outer diagonal, the central equations (5) dominate in number and thus require more computation. For this reason, as shown in this work, this section benefits most from the maximum achievable parallelism in order to achieve the best possible performance.

This work shows that data organization, the vectors storing the variables and constants of the algorithm in particular, should support higher memory bandwidth in the computing process for the most frequently accessed data. In particular, in order to increase memory bandwidth and the memory replication and adaptation of its structure were used, improving the performance of the implementation.

In order to achieve better performance, several additional optimizations were applied, including the replication of memories storing the most frequently accessed unknowns to alleviate bottlenecks, the parallelization of the accumulator and the preliminary computation of the floating-point divisions and storage of their results into a separate memory.

The goal of the optimization of the hardware solution within this work is the best possible performance considering the limitations imposed by the available resources on the target FPGA device.

The Xilinx Vitis HLS high-level synthesis tool was used for the hardware implementation. The software implementation was executed on a standard computer microprocessor.

Fig. 4 shows a representation of an accumulator. Although its implementation may differ, its functionality is identical to an adder, where the output is fed back to one of its inputs. Thus, its characteristics may be considered identical to those of a standard adder. In particular, for the implementation using the available digital signal processors (DSP), the typical case was considered with the initiation interval

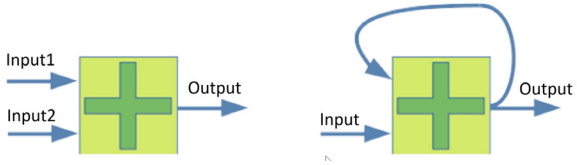


Fig. 4. Logical representation of an adder (left) and an accumulator (right)

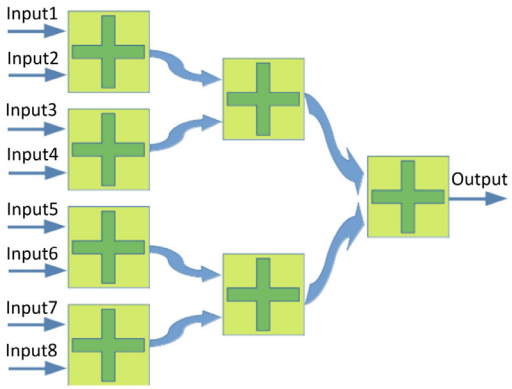


Fig. 5. Completely parallel accumulator

consisting of three cycles. Therefore, for SLE with N variables, the global delay for computing the error norm, required for enabling early completion of the algorithm, is three N cycles.

This computation time may become significantly longer than the actual computation of variables. For this reason, a parallel version of the accumulator is proposed. Fig. 5 shows the necessary structure of an accumulator for fully parallelizing the accumulation of eight values. It can be observed that the structure requires in total $N - 1$ adders and provides a result within $3*|\log_2 N|$ cycles. Therefore, while the performance improvement is considerable, the area requirements also increase substantially.

For this reason, employing a hybrid accumulator is advantageous, as shown in Fig. 6. A pipelined implementation of this approach allows to obtain a result with a speed of $3*(|\log_{2K}| + (N/K - 1))$ cycles, where K is the number of adders (i.e., the area) required by the implementation.

In this work, the accumulator was fully parametrized, and K becomes an arbitrary factor, determined by design choice. In this context, it is advisable to maximize K within the area constraints, determined by the maximum achievable parallelism of the remaining components of the accelerator.

For problems of large size and degrees of parallelism of several tens of computing units, the required memory bandwidth may exceed the capabilities of the platform, which, in our case, is limited to 460 GB/s. For this reason, efficient memory access becomes necessary.

Given the SLE under consideration, each instruction may involve up to five simultaneous accesses to the vector of variables (in the case of the computation of the error, while for the update of the variables up to four read accesses may be required at once).

Regarding the five-diagonal system under study, at each instruction two neighboring variables will be accessed, as well as two distant variables, corresponding to the outer diagonals located at a distance L from the central diagonal.

To the extent allowed by available resources, a window of relevant variables may be stored in internal registers, including neighboring and distant variables for all instructions that may execute in parallel. The variable from the left distant diagonal may be deleted from the registers when it is no longer needed in subsequent computations, releasing resources for new variables from the right distant diagonal.

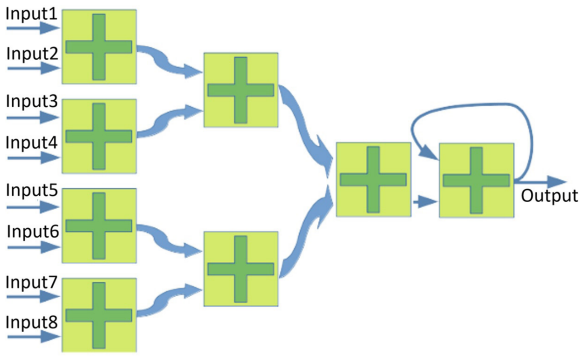


Fig. 6. Hybrid implementation of an accumulator

In this way, the variables may be stored in a cache in a sliding window fashion, so that memory access bandwidth to main memory is reduced: from five (or four) accesses to the memory of variables to one, achieving an approximate 30% reduction in required main memory bandwidth.

In some cases, the problem is so large and the distance L so great, that it's impossible to store all relevant variables in registers. In this case, even the storage of immediate neighbors leads to the reduction of memory bandwidth requirements by roughly 10%. A compromise solution is also possible, achieving an approximate 20% reduction.

Results and discussion

Research on the efficiency of the software and hardware implementations was carried out. The performance of the solution was taken as the criterion of efficiency when comparing results.

Considering that the achievable performance depends on the size of the problem, a number of matrices of different sizes was generated for this research.

Another important factor influencing the performance of the system, considering the iterative nature of the Jacobi method, is the precision required for the approximate solution. The more precise a result should be in order to be considered as acceptable, the more iterations should be executed, with a proportional increase in the processing time. In this work, values of *epsilon* (ε) within the range from 10^{-1} to 10^{-5} were investigated.

The results obtained for the software and hardware implementations show that the hardware implementation generally achieves a performance increase of approximately one order of magnitude compared to the software approach.

The system was described within the Vitis HLS of Xilinx environment[1], and the target device was the Alveo[2] [20] device: xcvu5p-fiva2104-1-e. Experimental results show that the time required to solve large SLE is reduced *by up to two orders of magnitude under given conditions*, compared to the standard software approach, based on a microprocessor.

Conclusion

The hardware implementation of SLE solver using the Jacobi iterative method, developed in this work, targets reconfigurable hardware accelerators and provides better performance than the software-based approach.

Using the resulting solution as a library component allows for increasing the efficiency of computing systems, achieving the best performance, when solving problems requiring the solution of SLE.

[1] Vitis HLS User Guide, Available: https://docs.amd.com/viewer/book-attachment/4lwvWeCi9jb~DWzdfWuVQQ/xRd0KUy2XBYs6mDFnJO6YA-4lwvWeCi9jb~DWzdfWuVQQ (Accessed 13.01.2026)

[2] Alveo U50 Data Center Accelerator Card Data Sheet: Alveo U50 Card Data Sheet (DS965), Available: https://docs.amd.com/r/en-US/ds965-u50 (Accessed 13.01.2026)

In the course of this work, a number of approaches to increasing the degree of parallelism were proposed, including parallelization and pipelining, implemented using high-level synthesis tools.

A further direction of the work is related to the analysis of the efficiency of hardware implementation for SLE solvers using other methods, considering their inherent parallelism.

REFERENCES

1. **Le Fèvre V., Herault T., Robert Y., Bouteiller A., Hori A., Bosilca G., Dongarra J.** Comparing the performance of rigid, moldable and grid-shaped applications on failure-prone HPC platforms. *Parallel Computing*, 2019, Vol. 85, Pp. 1–12. DOI: 10.1016/j.parco.2019.02.002

2. **Ashraf M.U., Alburaei Eassa F., Ahmad Albeshri A., Algarni A.** Performance and power efficient massive parallel computational model for HPC heterogeneous exascale systems. *IEEE Access*, 2018, Vol. 6, Pp. 23095–23107. DOI: 10.1109/ACCESS.2018.2823299

3. **Kalyaev I.A., Antonov A.P., Zaborovsky V.S.** Architecture of reconfigurable heterogeneous distributed supercomputer system for solving problems of intelligent data processing in the era of digital transformation of the economy. *Cybersecurity Issues*, 2019, Vol. 33, No. 5, Pp. 2–11. DOI: 10.21681/2311-3456-2019-5-02-11

4. **Dongarra J., Gottlieb S., Kramer W.T.C.** Race to exascale. *Computing in Science & Engineering*, 2019, Vol. 21, No. 1, Pp. 4–5. DOI: 10.1109/MCSE.2018.2882574

5. **Antonov A.P., Besedin D.S., Filippov A.S.** Research and comparative analysis of the effectiveness of software and hardware implementations of transposed matrix multiplication. *Computing, Telecommunications and Control*, 2024, Vol. 17, No. 1, Pp. 44–53. DOI: 10.18721/JCSTCS.17104

6. **Antonov A., Zaborovskij V., Kisilev I.** Developing a new generation of reconfigurable heterogeneous distributed high performance computing system. *Proceedings of International Scientific Conference on Telecommunications, Computing and Control*, 2021, Vol. 220, Pp. 255–265. DOI: 10.1007/978-981-33-6632-9_22

7. **Kastner R., Matai J., Neuendorffer S.** Parallel programming for FPGAs, *arXiv:1805.03648*, 2018. DOI: 10.48550/arXiv.1805.03648

8. **Foertsch J., Johnson J., Nagvajara P.** Jacobi load flow accelerator using FPGA. *Proceedings of the 37th Annual North American Power Symposium*, 2005, Pp. 448–454. DOI: 10.1109/NAPS.2005.1560554

9. **Tamuli M., Debnath S., Ray A., Majumdar S.** Implementation of Jacobi iterative solver in Verilog HDL, *2016 2nd International Conference on Control, Instrumentation, Energy & Communication* (*CIEC*), 2016, Pp. 103–105. DOI: 10.1109/CIEC.2016.7513747

10. **Morris G.R., McGruder R.Y., Abed K.H.** Accelerating a sparse matrix iterative solver using a high performance reconfigurable computer. *2010 DoD High Performance Computing Modernization Program Users Group Conference*, 2010, Pp. 517–523. DOI: 10.1109/HPCMP-UGC.2010.30

11. **Ruan H., Huang X., Fu H., Yang G.** Jacobi solver: A fast FPGA-based engine system for Jacobi method. *Research Journal of Applied Sciences, Engineering and Technology*, 2013, Vol. 23, No. 6, Pp. 4459–4463. DOI: 10.19026/rjaset.6.3452

12. **Pelipets A.V.** Rasparallelivanie iteratsionnykh metodov resheniia sistem lineinykh algebraicheskikh uravnenii na rekonfiguriruemykh vychislitel'nykh sistemakh [Parallelization of iterative methods for solving system linear algebraic methods on reconfigurable computing resources]. *Superkomp'iuternye Tekhnologii* (*SKT-2016*) [*Supercomputer Technologies*], 2016, Vol. 1, Pp. 194–198.

13. **Chekina M.D.** Modification of the implementation of the Jacobi method in simulating superdiffusion of radon on reconfigurable computer systems. *Izvestiya SFedU. Engineering Sciences*, 2021, Vol. 7, Pp. 198–206. DOI: 10.18522/2311-3103-2021-7-198-206

14. **Levin I.I., Dordopulo A.I., Pelipets A.V.** Implementation of iteration methods for solution of linear equation systems in problems of mathematical physics on reconfigurable computer systems. *Bulletin of the*

South Ural State University. Series: Computational Mathematics and Software Engineering, 2016, Vol. 5, No. 4, Pp. 5–18. DOI: 10.14529/cmse160401

15. **Levin I.I., Dordopulo A.I., Pisarenko I.V., Melnikov A.K.** Description of Jacobi algorithm for solution of linear equation system in architecture-independent set@l programming language. *Izvestiya SFedU. Engineering Sciences*, 2018, Vol. 5, Pp. 34–48. DOI: 10.23683/2311-3103-2018-5-34-48

16. **Morris G.R., Abed K.H.** Mapping a Jacobi iterative solver onto a high-performance heterogeneous computer. *IEEE Transactions on Parallel and Distributed Systems*, 2013, Vol. 24, No. 1, Pp. 85–91. DOI: 10.1109/TPDS.2012.121

17. **Pourhaj P., Teng D.H.-Y., Wahid K., Ko S.-B.** System size independent architecture for Jacobi processor. *2008 Canadian Conference on Electrical and Computer Engineering*, 2008, Pp. 002033–002036. DOI: 10.1109/CCECE.2008.4564902

18. **Uguen Y., Dinechin F.** Design-space exploration for the Kulisch accumulator, 2017, Available: https://hal.science/hal-01488916/file/kulisch-acc-2017.pdf (Accessed 23.12.2025)

INFORMATION ABOUT AUTHORS / СВЕДЕНИЯ ОБ АВТОРАХ

Mauricio F. Gonzalez
Гонзалез Мауризио Фаюла
E-mail: fayula.gm@edu.spbstu.ru

Alexander P. Antonov
Антонов Александр Петрович
E-mail: antonov_ap@spbstu.ru
ORCID: https://orcid.org/0000-0002-4107-8950

Submitted: 27.05.2025; Approved: 05.12.2025; Accepted: 25.12.2025.
Поступила: 27.05.2025; Одобрена: 05.12.2025; Принята: 25.12.2025.

Computing, Telecommunication and Control, 2025, Vol. 18, No. 4, Pp. 87–101.
Информатика, телекоммуникации и управление. 2025. Том 18, № 4. С. 87–101.

Overview article
DOI: https://doi.org/10.18721/JCSTCS.18408
UDC 004.9:004.94

MITIGATING DATA GROWTH IN POW BLOCKCHAINS: STORAGE REDUCTION METHODS FOR SCALABILITY WITHOUT COMPROMISING DECENTRALIZATION

D.D. Razuvaev[1,2] ✉ , *S.M. Ustinov*[1]

[1] Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russian Federation;

[2] Lomonosov Moscow State University, Moscow, Russian Federation

✉ Razuvaev_DD@mail.ru

Abstract. The exponential growth of data volume in Proof-of-Work (PoW) blockchains threatens their decentralization. The article provides a systematic analysis of data growth mitigation methods (sharding, block pruning, off-chain storage etc.), identifying their key flaws: compromised auditability, increased synchronization complexity or centralization. To address the issue, a novel Periodic Aggregation with Dual Hash Anchoring (PADHA) method is proposed. Its key innovation is the synergy of data pruning and the controlled use of chameleon hash functions. The method enables linear reduction of stored history by creating final state aggregators and subsequent secure "cleansing" of blocks from past epochs. PADHA preserves cryptographic chain integrity and PoW support without trusted third parties. The method is designed for Fact-Oriented Blockchains that store final data states (facts), making it promising for registries, IoT and other applications where current information, not its change history, is critical.

Keywords: distributed registries, decentralized systems, blockchain, scalability, decentralization, block pruning

Citation: Razuvaev D.D., Ustinov S.M. Mitigating data growth in PoW blockchains: Storage reduction methods for scalability without compromising decentralization. Computing, Telecommunications and Control, 2025, Vol. 18, No. 4, Pp. 87–101. DOI: 10.18721/JCSTCS.18408

Обзорная статья
DOI: https://doi.org/10.18721/JCSTCS.18408
УДК 004.9:004.94

ОПТИМИЗАЦИЯ ХРАНЕНИЯ ДАННЫХ В PROOF-OF-WORK БЛОКЧЕЙН-СИСТЕМАХ: МЕТОДЫ СОКРАЩЕНИЯ ОБЪЕМА И ОБЕСПЕЧЕНИЕ МАСШТАБИРУЕМОСТИ В УСЛОВИЯХ ДЕЦЕНТРАЛИЗАЦИИ

Д.Д. Разуваев[1,2] ✉ , *С.М. Устинов*[1]

[1] Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Российская Федерация;

[2] Московский государственный университет им. М.В. Ломоносова, Москва, Российская Федерация

✉ Razuvaev_DD@mail.ru

Аннотация. Экспоненциальный рост объема данных в блокчейнах на Proof-of-Work (PoW) угрожает их децентрализации. В статье системно анализируются методы борьбы с ростом данных (шардинг, обрезка блоков, офф-чейн хранение и др.) и выявляются их ключевые недостатки: нарушение аудируемости, усложнение синхронизации или централизации. Для решения проблемы предлагается новый метод периодической агрегации с двойным хеш-якорением (PADHA). Его ключевая инновация – синергия обрезки данных и контролируемого использования хеш-функций-хамелеонов. Метод обеспечивает линейное сокращение хранимой истории за счет создания агрегаторов финального состояния и последующего безопасного «очищения» блоков прошедших эпох. PADHA сохраняет криптографическую целостность цепочки и поддержку PoW, не требуя доверенных третьих сторон. Метод применим для факт-ориентированных блокчейнов (FOB), хранящих итоговые состояния данных (факты), что делает его перспективным для реестров, IoT и других приложений, где критична актуальная информация, а не история ее изменений.

Ключевые слова: распределенные реестры, децентрализованные системы, блокчейн, масштабируемость, децентрализация, обрезка блоков

Для цитирования: Razuvaev D.D., Ustinov S.M. Mitigating data growth in PoW blockchains: Storage reduction methods for scalability without compromising decentralization // Computing, Telecommunications and Control. 2025. Т. 18, № 4. С. 87–101. DOI: 10.18721/JCSTCS.18408

Introduction

Currently, blockchain technology is widely applied across various domains including finance [1, 4], intellectual property management [2], insurance, tourism [3], healthcare and biomedical systems [10, 21], government services and education [5, 9], as well as 5G networks [7, 8], to address challenges related to data integrity, transparency and disintermediation. Blockchain is a distributed digital registry based on the principles of decentralization, cryptography and consensus [11, 25]. Unlike traditional centrally managed databases, blockchain systems function as peer-to-peer networks of nodes (participants), each of which usually stores a complete copy of information. The data in such a system is organized as a sequence of blocks [22], where each block contains:

- a set of records (for example, transactions or any other data);
- hash identifier – a unique cryptographic signature calculated based on the block content;
- hash of the previous block, which ensures a cryptographic link between the chain elements.

This architecture guarantees the immutability of data (excluding the use of hash chameleons): any attempt to change information in existing block will disrupt communication with subsequent blocks,

which will be immediately detected by the network. To add new blocks, a consensus algorithm [24, 26] is used – a set of rules that allows nodes to coordinate the state of the system without trusting a central authority. The first implementation of this technology, which appeared in 2008, demonstrated the potential of blockchain technology and established itself as a tool for creating censorship- and fraud-resistant systems, as well as revealed fundamental limitations associated with the scalability of such systems.

A key feature of the Proof-of-Work (PoW) blockchain [27] is the mechanism for achieving consensus through computational tasks [28]. Nodes (miners) compete to solve a cryptographic puzzle that requires significant resources. The winner gets the right to add a block to the chain and a reward, and the other nodes check the correctness of the decision. This process, although it provides a high level of security, leads to an increase in data, since each new block increases the total volume of the chain, while deleting a block violates the integrity of the blockchain.

Over time, this creates "a paradox of centralization of decentralized systems": demands on computing power and storage become so high that participation in the network becomes available only to a limited number of specialized participants (who can commercialize their work by providing Software-as-a-Service solutions). For example, in January 2025, the size of the blockchain in the first implementation of the technology exceeded 600 GB, and in March 2025 it already stands at 612 GB[1], and the annual energy consumption of the network is comparable to that of entire countries [6]. These factors not only threaten the stability of the system, but also limit its application in areas where data processing speed and energy efficiency are critical, such as IoT, mobile devices and systems with data processing speed requirements. Fig. 1 shows a graph of the size of the data volume over a 10-year time scale, which clearly demonstrates the problem of data accumulation.

Thus, despite its revolutionary potential, PoW-based blockchain systems face a trilemma: scalability, decentralization and security cannot be simultaneously optimized within a classical architecture. Resolving this problem requires rethinking approaches to data storage and management, which is the focus of this study.

The exponential increase in data volume in PoW-based blockchain systems is a direct consequence of their architectural features. Each new block added to the chain not only expands the data history, but also requires all network participants to constantly verify and store a complete copy of the registry. This leads to a number of systemic contradictions:

• Accumulation of "historical load"

In classic PoW implementations [23] (for example, in bitcoin), the size of the chain increases by 1–4 MB daily (this value is individual for each implementation), which over 5 years of operation has created a load of several hundred gigabytes. For nodes with limited resources (for example, mobile devices), this makes participation in the network technically and economically impractical.

• Energy-computing imbalance

The PoW mechanism requires repeated recalculation of hashes to achieve consensus, which leads to duplication of computing operations on all nodes [18]. The growing volume of data exacerbates this problem: the validation of a long chain of blocks consumes more and more energy.

• Degradation of network synchronization

Increasing the download and verification time of the blockchain reduces the synchronization speed of new nodes. In high-bandwidth networks (for example, Ethereum before switching to Proof-of-Stake), full synchronization can take days, increasing the risks of chain splits and reducing resistance to attacks.

• Limitation of functionality of light clients

"Simplified" nodes (light clients) that do not store a complete copy of the blockchain are forced to rely on trusted third-party services to access the data. This violates the principle of decentralization and creates vulnerabilities such as censorship or information substitution.

[1] Protocol Labs. Filecoin: A decentralized storage network, 2017. Available: https://filecoin.io/filecoin.pdf (Accessed 10.05.2025)

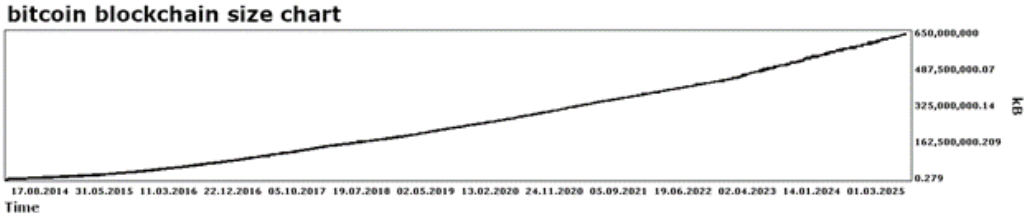


Fig. 1. Graph of the size of the Bitcoin blockchain over time

Attempts to solve these problems through traditional methods – sharding, block pruning or hybrid consensus – face fundamental limitations. For example, sharding, which divides the network into sub-chains, reduces the load on individual nodes, but increases the complexity of cross-shard operations and the risk of conducting attacks in isolated segments (shards) when conducting attacks of the "51 percent" class, where an attacker with more than 50% of the capacity (>50% of the total hashrate, which reflects how many cryptographic operations to find a block solution in the blockchain can be performed by devices in a unit of time), can single-handedly control a sub-chain or the entire block chain. Pruning, although it reduces the local amount of data, deprives the network of the total PoW power.

These contradictions actualize the search for alternative approaches that will preserve the advantages of PoW (for example, resistance to Sybil attacks), and eliminate the problem of endless growth of the stored amount of data. The key direction is the development of protocols that optimize information storage without compromising security, for example, through segmentation of data by relevance level or the introduction of mechanisms for "forgetting" outdated blocks.

Relevance of the topic

The relevance of this study is due to the systemic crisis of scalability faced by PoW blockchain solutions in the context of exponential growth of stored data. Despite the widespread introduction of technology into finance [1], intellectual property [2], logistics, healthcare [21] (including use in biomedical systems [10]), tourism and the public sector [3, 5, 9], blockchain technology has both advantages and fundamental limitations [13] – immutability data and full replication are becoming a barrier to sustainable development. Traditional optimization methods such as sharding or hybrid consensus demonstrate partial efficiency, but do not solve the key problem: inconsistencies between the growing volume of data and the requirements of decentralization[2] [12].

The scientific significance of this work lies in addressing several key gaps.

First, there is a lack of comprehensive analysis regarding how existing data reduction methods – such as multi-chain networks, off-chain storage and modifiable blockchains[3] [14–16] – affect the fundamental blockchain trilemma of security, decentralization and scalability.

Second, the theoretical framework lacks well-defined criteria for assessing "historical redundancy" of data; for instance, clear principles are needed to determine which blocks can be safely deleted or compressed without undermining the system's auditability.

Furthermore, a pragmatic paradox emerges: technologies originally designed to eliminate centralized intermediaries, such as cryptocurrencies [4], increasingly depend on centralized cloud storage solutions to archive old blocks [14], which fundamentally contradicts their decentralized ideology.

The practical relevance of this research is driven by pressing industry demands. Financial institutions seek to reduce the operational costs of maintaining full PoW network nodes while adhering to stringent regulatory standards like the European General Data Protection Regulation (GDPR).

[2] Huobi Research Institute Report. Game of Thrones in Blockchain: Multi-Chain Networks Battle for Supremacy, 2022.
[3] Huobi Research Institute Report. Game of Thrones in Blockchain: Multi-Chain Networks Battle for Supremacy, 2022.

Simultaneously, IoT [15] and 5G networks [7] require lightweight blockchain solutions capable of real-time data processing. Additionally, state registries – for land, education [9] and other public records – confront the challenge of long-term data preservation due to bandwidth constraints and ever-growing data volumes.

Therefore, this study conducts a systematic comparative analysis of data management methods in PoW systems, including:

- multichain architectures (AppChain)[4];
- off-chain storage with lazy loading [14];
- protocol modifications via hash chameleon functions [16].

The conducted analysis reveals the fundamental trade-offs inherent to each of the considered approaches. For instance, aggressive block pruning leads to the irreversible loss of the complete history and, consequently, to the ability to perform an independent chain audit. Hybrid consensus models, while aiming to reduce energy consumption, carry inherent risks of power centralization among a limited circle of stakeholders. Finally, modifiable blockchains based on chameleon hashes, while addressing data growth, introduce vulnerabilities related to the potential for targeted collision attacks and the centralization of editing control, which undermines the core principle of immutability. These identified limitations necessitate fundamentally new solutions based on the following design principles: dynamic data lifecycle management, enabling the deletion or archiving of obsolete information; segmentation of data based on its criticality to system operation; and minimizing reliance on external decentralized repositories (such as IPFS or Storj) to preserve the blockchain's self-sufficiency and security.

To overcome the noted shortcomings of existing methods, this article, based on their systematic analysis, proposes a new method of Periodic Aggregation with Dual Hash Anchoring (PADHA). Its goal is to resolve the contradiction between the need to preserve the key advantages of PoW, such as resistance to attacks and censorship, and the Web 3.0 requirements for high data processing speed and energy efficiency. The results of the study form the basis for the further development of data storage optimization methods[5] [19].

Methods

To combat the exponential increase in data volume in PoW systems, various approaches have been proposed that can be classified according to the level of impact on the blockchain architecture.

The effectiveness of data growth control methods in PoW systems directly depend on their ability to maintain a balance between volume optimization, decentralization, and security. However, most solutions upset this balance by introducing unacceptable trade-offs, especially in the context of PoW, where the integrity of the chain and the availability of the full history are the basis of consensus.

Traditional methods

Sharding

Bottom line: dividing the network into independent sub-chains (shards), each of which processes part of the transactions[6].

Example: Zilliqa, Ethereum before switching to Proof-of-Stake (PoS).

Advantages: reducing the load on individual nodes, increasing throughput.

Disadvantages: increased complexity of cross-shard operations; increased vulnerability of sub-chains to a sub-chain attack in cases where more than 50% of the hashrate of the sub-chain accumulates in the attacker; violation of the integrity of data auditing.

[4] Huobi Research Institute Report. Game of Thrones in Blockchain: Multi-Chain Networks Battle for Supremacy, 2022.
[5] Buterin V. Merkling in Ethereum, 2015. Available: https://blog.ethereum.org/2015/11/15/merkling-in-ethereum (Accessed 10.05.2025)
[6] Huobi Research Institute Report. Game of Thrones in Blockchain: Multi-Chain Networks Battle for Supremacy, 2022.

Key limitation: splitting into sub-chains complicates synchronization and increases the risk of attacks in small shards. This method complicates synchronization and increases the risk of attacks aimed at destabilizing the network.

Pruning

Bottom line: deleting outdated blocks while preserving only headers or critical data (for example, snapshots of the Unspent Transaction Output status) [12].

Example: Bitcoin Core ("pruned node" mode).

Advantages: reduction of the local data volume to 5–10% of the original.

Disadvantages: loss of the possibility of independent verification of the complete history; loss of historical validity; dependence on complete archive nodes, which leads to centralization.

Key limitation: deleting the initial blocks (including the genesis block) violates the integrity of the hash chain, which contradicts the PoW principle. This method may disrupt the stability of the network.

Hybrid consensus models (*PoW/PoS*)

Bottom line: a combination of PoW for creating blocks and PoS for validation[7].

Example: Decred [16].

Advantages: reduction of energy consumption and performance requirements due to partial abandonment of mining.

Disadvantages: conflicts between consensus mechanisms; vulnerability to "nothing-at-stake" attacks; risk of centralization of power among stakeholders.

Key limitation: the combination of PoW and PoS creates contradictions between miners and stakeholders, and also leads to the centralization of power among large stakeholders and a potential decrease in the security of the blockchain.

Innovative methods

Multi-chain networks

Bottom line: creating a hierarchy of blockchains, where the main chain (Layer 1) coordinates the work of sidechains (Layer 2)[8].

Example: Polkadot (based on Nominated PoS) [17], Cosmos (based on Tendermint BFT), Rootstock RSK (PoW).

Advantages: isolation of application data; scalability due to parallelism.

Disadvantages: the difficulty of synchronization between circuits.

Key limitation: the hierarchy of blockchains complicates auditing and synchronization [20]. Synchronization problems directly reduce the stability of blockchain systems.

Off-chain storage with lazy loading

Bottom line: transferring old blocks to external decentralized storage (for example, InterPlanetary File System[9]) with on-demand download [14].

Example: Arweave (permanent storage), Filecoin (archiving on request)

Advantages: reducing the load on the nodes; maintaining access to the full history.

Disadvantages: delays in requesting archived data; vulnerability to storage failures; dependence on the stability of third-party networks.

Key limitation: transferring data to IPFS/Filecoin introduces delays and risks of information loss, and does not solve the problem of blockchain bloat, shifting responsibility for the preservation of archived data to a distributed solution. The self-sufficiency of the blockchain is being violated.

Redactable blockchains

Bottom line: using hash chameleon functions to edit or delete blocks without violating the integrity of the chain [16].

[7] Huobi Research Institute Report. Game of Thrones in Blockchain: Multi-Chain Networks Battle for Supremacy, 2022.
[8] Huobi Research Institute Report. Game of Thrones in Blockchain: Multi-Chain Networks Battle for Supremacy, 2022.
[9] Protocol Labs. Filecoin: A decentralized storage network, 2017. Available: https://filecoin.io/filecoin.pdf (Accessed 10.05.2025)

Example: Hyperledger Fabric (primarily in enterprise contexts)[10].

Advantages: dynamic data management.

Disadvantages: centralization of control (modifiers); vulnerability to collision attacks.

Key limitation: editing blocks using hash chameleon functions undermines the basic principle of block immutability. For systems that prioritize the principle of immutability of information, this method is not optimal. The risk of collision attacks harms the data integrity paradigm in the system.

Merkle Patricia Trie

Bottom line: optimizing the storage of network status through tree-like hash structures that allow you to remove duplicates [8].

Example: Ethereum State Trie

Advantages: reduction of data volume by 30–50%; acceleration of transaction search.

Disadvantages: the complexity of the implementation for PoW networks; the risk of data loss in case of failures.

Key limitation: this method is applicable to the classical implementation of the PoW blockchain only to a limited extent, for working with the state tree stored inside the blockchain. At the same time, the method solves the problem of inflating the PoW of the blockchain system by using pruning method, with all the conclusions drawn from this.

Summarizing the review and analysis of existing solutions, addressing the inherent limitations of existing methods is challenging. However, a viable path forward involves introducing specific architectural constraints. These constraints narrow the method's applicability to a particular class of tasks – specifically, systems that only need to store final state data rather than complete transactional histories. While this represents a focused application domain, it remains broad enough to cover numerous practical use cases, such as registries, sensor data logging and state tracking. The significant gain achieved – dramatic data volume reduction without breaking cryptographic chain integrity – justifies this targeted approach.

Proposed Method

An analysis of existing traditional and innovative methods for combating data growth in PoW blockchain systems has revealed fundamental problems affecting the scalability, security and/or decentralization of solutions based on them. Models of sharding, block pruning and hybrid consensus models, while reducing the load on nodes, violate the principles of data integrity, auditability and participant equality. Architectures with application-specific sidechains and off-chain storage, in turn, introduce dependency of the main chain on third-party networks and complicate synchronization or lead to forced centralization of solutions. Furthermore, modifiable blockchains based on chameleon hashes, despite offering dynamic data management, also contain inherent risks.

For an unambiguous definition of the proposed method and its applicability boundaries, it is necessary to introduce the following terminology:

• Atomic Fact – an immutable, self-contained unit of information representing an arbitrary set of data (an assertion), which does not require references to other facts for its interpretation or integrity.

• Fact Block – an operationally atomic structural unit of a distributed ledger, whose sole content is an ordered set of atomic facts, supplemented by a unified service header (timestamp, hash). The acceptance or rejection of a fact block by the system follows an "all-or-nothing" principle.

• Fact-Oriented Blockchain (FOB) – a distributed ledger implemented as a chronologically and cryptographically linked chain of fact blocks, which guarantees the immutability and order of atomic facts.

This method is specifically applicable to FOB, which imposes a key architectural constraint: such a system can only store final data (facts), but cannot store procedural or transactional records describing the process of changes.

[10] Hyperledger Fabric. A Blockchain Platform for the Enterprise, 2021. Available: https://hyperledger-fabric.readthedocs.io/en/release-2.5/ (Accessed 10.05.2025)

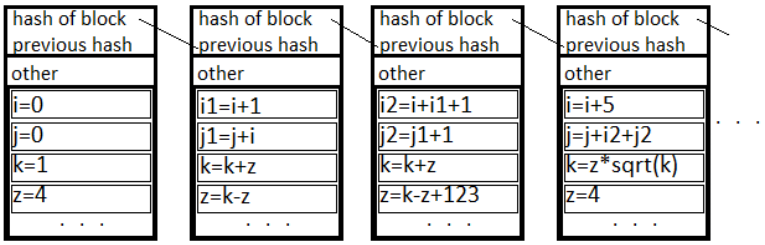


Fig. 2. Blockchain with transactions

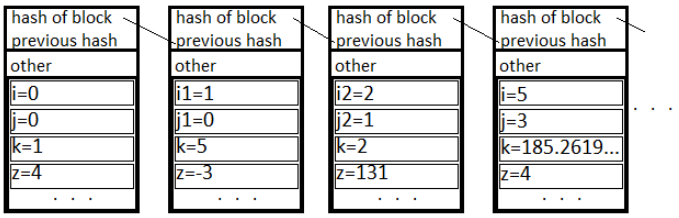


Fig. 3. Fact-Oriented Blockchain

The main difference between a classical blockchain model and an FOB lies in the nature, connectivity and interpretation of the data stored in the blocks (Figs. 2, 3).

In a classical blockchain (Fig. 2), blocks typically contain transactions or change commands (e.g., i = i + 5). The information in subsequent blocks is semantically and logically dependent on the information in previous blocks. To obtain the current state of the system (for example, the final value of *j*), it is necessary to replay the entire transaction history from the very beginning (the genesis block). The deletion or corruption of an intermediate block makes it impossible to calculate or verify any subsequent data.

In FOB (Fig. 3), this limitation is eliminated due to a different organization of data. Each atomic fact within a block represents a final value or statement (for example, i = 3), which is semantically complete and does not require referencing other facts for its interpretation. Let us define that a system based on FOB interprets data by treating the last value recorded in the chain for a given entity (for example, variable i) as its current state. Thus, a new fact i = 5 in a later block does not reference the previous value but semantically overwrites it for the system. A fact-block serves for the operationally atomic batch transfer of such independent facts.

Thus, a situation is achieved where:

- information (an atomic fact) in any block is self-sufficient for interpretation;
- understanding the current state of the system does not require reproducing the full history of changes; it is sufficient to read the latest facts for the entities of interest;
- integrity of a block is ensured cryptographically, and the semantic value of each fact is contained within itself and its position in the chain, which determines the relevance of the value.

To solve the problem of uncontrolled “bloating”, the use of a combination of two method concepts is proposed: the pruning method and the editable blockchains method.

The classic implementation of the data pruning method involves removing transactions that are technically unnecessary for the target user. This results in the loss of the entire chain’s validity and reduces the weight of the blockchain (Fig. 4).

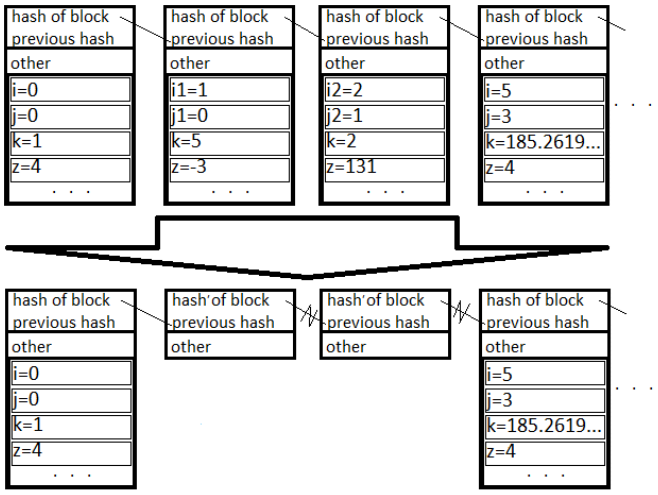


Fig. 4. Pruning method for an FOB: in blocks 2 and 3,
the facts have been removed and the hash chain is broken

The second method proposes using a chameleon hash function instead of a regular hash function. A chameleon hash function allows generating collisions using a special key. Fig. 5 demonstrates the method of using an editable blockchain based on chameleon hash functions.

The PADHA method is designed for FOB and combines the advantages of data pruning and the use of chameleon hash functions, enabling a radical reduction in the volume of stored data without losing the cryptographic integrity of the chain. The core idea of the method is the periodic creation of special aggregator blocks, which contain the final (current) values of all entities for a certain period, and the subsequent "clearing" (replacing the bodies with empty ones) of ordinary blocks in that period while preserving the hash chain thanks to chameleon hash functions.

Architectural features and prerequisites

As noted, an FOB operates with atomic facts, each of which is a self-sufficient statement (e.g., "i = 5"). The system interprets the most recent (in chain order) value for a given entity (i) as the current one. This property enables state aggregation: instead of storing all intermediate changes, it is sufficient to store only the latest values at the moment of aggregation.

In the PADHA method, each block contains two cryptographic hashes: a regular one (e.g., SHA-256) and a chameleon hash. The link between blocks is considered valid until an aggregator block is created if at least one of the two hashes matches (i.e., either the regular hash or the chameleon hash), or the link between blocks is valid if both hashes match when the aggregator block has not yet been formed (the epoch is not yet completed). Such link validation rules allow, after cleaning the block bodies, the chain to remain valid via chameleon hashes, while the regular hashes may become invalid due to changes in the block contents. They also protect against malicious attempts to modify blocks during an unclosed epoch.

Method description

1. Normal operation mode (within an epoch)

The blockchain is divided into epochs of a fixed length (for example, 1000 blocks). During an epoch, all blocks are created in the normal mode: each block contains atomic facts (changes in entity values) and has a header with two hashes: hash_std (standard hash) and hash_cham (chameleon hash). Both hashes are calculated based on the block's content and the corresponding hash of the previous block. Thus, at the beginning of an epoch, the chain is valid according to both hashes (Fig. 6).

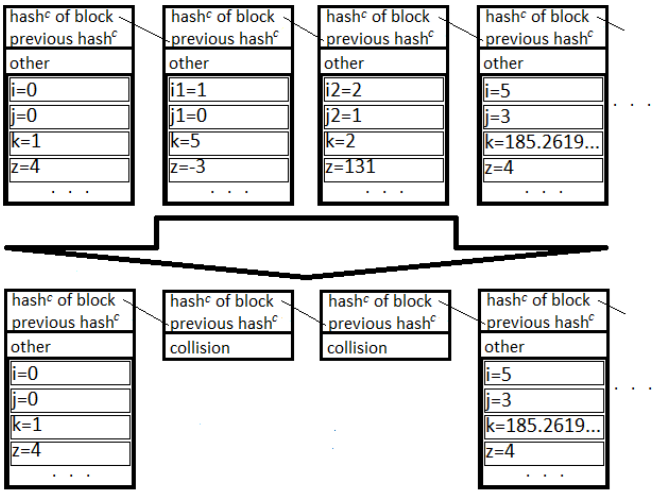


Fig. 5. Chameleon hash method: blocks 2 and 3 have been cleared,
the hash chain remains intact because collisions have been inserted

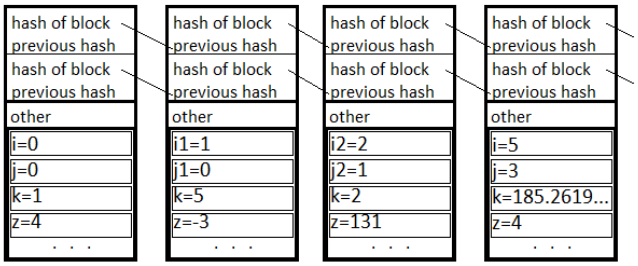


Fig. 6. PADHA – operation within an epoch with a length of 5

2. Creating an aggregator block

At the end of each epoch (every 1000th block), a special aggregator block is created. This block does not contain ordinary facts; instead, it contains:

• The final values of all entities that were changed during the epoch (essentially, the last assignments for each variable).

• The aggregator may also contain values of entities that were not changed during the epoch but are relevant at the time of aggregation (i.e., a complete snapshot of the system state). However, considering that an FOB can contain a vast number of entities, it is more practical to include only those entities that were changed in the given epoch. In this case, to obtain the current value of any entity, it will be necessary to find the last aggregator in which it was changed or read the value from the current epoch (if it has not yet been aggregated).

The aggregator also has two hashes that reference the corresponding hashes of the previous (999^{th}) block of the epoch (Fig. 7).

3. Cleaning of epoch blocks

After the aggregator is created and accepted by the network, the process of cleaning the blocks of this epoch (from the 1^{st} to the 999^{th}) begins. Cleaning consists of replacing the body of each block with a collision (essentially clearing all atomic facts in the block) in such a way that the block's chameleon hash remains unchanged. This is achieved due to the property of the chameleon hash function:

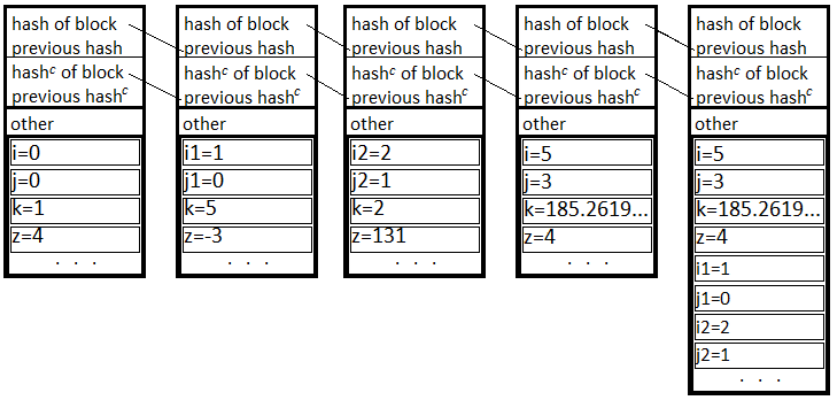


Fig. 7. PADHA – generation of an epoch aggregator block with a length of 5

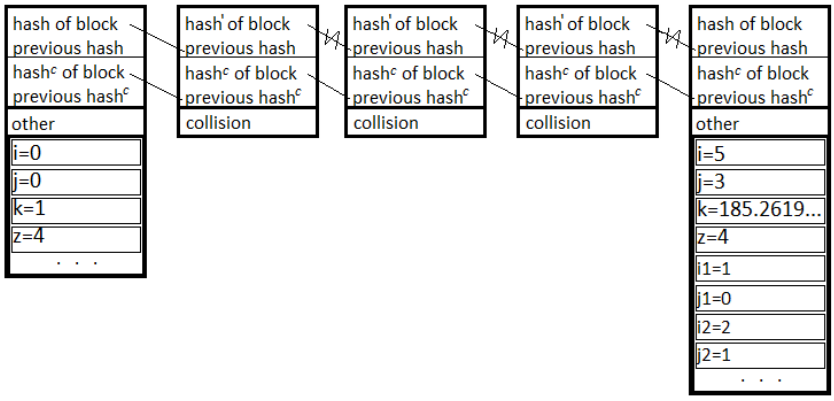


Fig. 8. PADHA – data deletion, chameleon hash correction, completion of an epoch with a length of 5

knowing a special key (which is publicly available in the system), one can find a new body such that the chameleon hash matches the original one. In this case, the regular hash of the block will change.

After cleanup, the connection between blocks within an epoch remains valid only via hash chameleons. The standard hashes no longer form a continuous chain. However, since we allow validation of blocks up to the aggregator using either of the two hashes, the chain remains valid.

4. Transition to the next epoch

After the epoch blocks are cleared, the new blocks (starting from block 1001) reference the aggregator (block 1000) and subsequent blocks accordingly via both hashes. Thus, the aggregator becomes the new reference point for the next epoch. It is important to note that the aggregator itself is not cleared (it remains full), as it contains the summary information necessary for state recovery.

5. Multiple aggregation and history compression

The process repeats every epoch. After creating the aggregator for epoch 2 (block 2000) and cleaning blocks 1001–1999, we are left with a chain consisting of:

- genesis block (full);
- cleaned blocks of epoch 1 (1–999);
- aggregator of epoch 1 (block 1000, full);
- cleaned blocks of epoch 2 (1001–1999);
- aggregator of epoch 2 (block 2000, full), and so on.

6. Recursive aggregation

To further reduce the data volume, recursive aggregation can be applied: after creating the aggregator for epoch 2, which also includes information from aggregator 1, the aggregator of epoch 1 can also be cleaned (replaced with an empty one while preserving its chameleon hash), because the final state of all entities by the end of epoch 2 will be contained in the aggregator of epoch 2.

Advantages of the PADHA method:

• Significant reduction of data volume: The bodies of blocks, except for the last aggregator, can be cleared. This provides a linear (or even logarithmic with recursive aggregation) reduction in the volume of stored history.

• Preservation of cryptographic integrity: Thanks to the use of chameleon hash functions, the block chain remains cryptographically linked, and chain validation is possible (via chameleon hashes).

• Support for PoW: Since block headers (including the chameleon hash) remain unchanged, the proof of work performed for each block remains valid. Mining new blocks is not disrupted.

• Decentralization: The method does not require trusted third parties or centralized archives. All nodes can independently perform the clearing, as the chameleon hash key is public. The chameleon key is public for all participants because, prior to epoch closure, blocks at the protocol level must maintain validity through both chameleon hash and standard hash links, and validation is performed against both hashes. Consequently, any unauthorized modification of a chameleon hash and its corresponding block body will disrupt the standard hash chain and be rejected by the network. After an epoch is closed, the blocks are "cleansed" (their bodies are emptied) at the protocol level, meaning they cannot legitimately contain data. Should data appear in such blocks, the network will also reject them.

• A malicious actor cannot create an alternative history without performing PoW for all blocks.

Limitations of the PADHA method:

• Requirement for a fact-oriented model: The method is only applicable to blockchains where data is represented as atomic facts, and the current state is determined by the last value. It is not suitable for transactional models where history is important.

Conclusion

In response to the limitations of existing methods, this paper proposes a novel method of PADHA, designed for FOBs. The exponential growth of data volume in PoW blockchain systems represents a fundamental challenge to their long-term sustainability and decentralization. The systematic analysis conducted in the article confirmed that traditional optimization methods – sharding, pruning of outdated blocks and hybrid consensus models – do not eliminate the key contradiction between the need to reduce "historical load" and preserve the basic properties of PoW: immutability, full chain auditability and distributed consensus. Each of these approaches introduces unacceptable compromises, whether it is the complication of synchronization and increased risks of attacks (sharding), the loss of the possibility for independent audit of the full history (pruning) or the conflict of consensus mechanisms (hybrid models). The PADHA method key innovation lies in the synergy of pruning principles and the controlled use of chameleon hash functions, ensures linear reduction in the volume of stored data through the periodic creation of aggregator blocks containing the system's final states, followed by cryptographically secure "cleansing" of the bodies of blocks from past epochs. This achieves the preservation of cryptographic chain integrity through dual hash anchoring and full support for the PoW mechanism, which distinguishes this approach from classical pruning or editable blockchains. However, the method's application necessitates introducing specific architectural constraints, narrowing its applicability to systems designed to store only final state data (facts) rather than complete transactional histories. While this represents a focused application domain, it remains broad enough to cover numerous practical use cases, such as registries, sensor data logging and state tracking. The significant gain achieved – dramatic data volume reduction without breaking cryptographic chain integrity – justifies this targeted approach.

Compared to existing methods, the proposed PADHA approach offers distinct qualitative advantages. Unlike sharding, it maintains a single coherent chain without cross-shard complexity. In contrast to classical pruning, it preserves cryptographic chain integrity via chameleon hashes. And unlike editable blockchains, it employs a controlled, protocol-level use of chameleon functions that does not undermine immutability for data. This enables the creation of scalable and energy-efficient PoW systems suitable for deployment in resource-constrained environments (IoT, mobile devices), without compromising their decentralization and security. The directions for further research are: the development of cryptographic proofs of the aggregators' correctness to minimize trust, the creation of adaptive protocols for selecting compression parameters, and an in-depth analysis of the method's resilience to new attack vectors.

Thus, the presented PADHA method offers a concrete path to overcoming the key limitation of PoW blockchains, demonstrating that data storage optimization is achievable not through the abandonment of fundamental principles, but through their adaptive evolution and the application of modern cryptographic primitives. This opens prospects for a new generation of decentralized systems that combine the robustness of PoW with the scalability demands of the Web 3.0 environment.

REFERENCES

1. **Yusupova D.R., Tazetdinova L.R.** The advantages of using blockchain technology in the insurance business. *Forum molodykh uchenykh* [*Forum of Young Scientists*], 2019, Vol. 33, No. 5, Pp. 1437–1442.

2. **Shomakhov A.R.** Use of blockchain technologies in the field of intellectual property rights. *Voprosy studencheskoi nauki* [*Student Science Issues*], 2020, Vol. 51, No. 11, Pp. 134–137.

3. **Marchenko K.Iu., Gimadeeva A.S.** Vnedrenie tekhnologii blokchein na mezhdunarodnyi rynok turistskikh uslug [Implementation of blockchain technology in the international tourism services market]. *Vestnik sovremennykh issledovanii* [*Bulletin of modern research*], 2018, Vol. 21, No. 6.4, Pp. 231–234.

4. **Kozin A.D., Makarova L.N., Buter A.P.** Realizatsiia tekhnologii blokchein na primere kriptovaliut [Implementation of blockchain technology using cryptocurrencies]. *Evraziiskii Soiuz Uchenykh* [*Eurasian Union of Scientists*], 2020, Vol. 81, No. 12, Pp. 10–17.

5. **Agamalian N.Kh.** Blokchein v auditorskoi deiatel'nosti [Blockchain in auditing activities]. *E-Scio*, 2020, Vol. 46, No. 7, Pp. 449–455.

6. **Rîndaşu S.-M.** Blockchain in accounting: Trick or treat? *Quality – Access to Success*, 2019, Vol. 20, Pp. 143–147.

7. **Bezzateev S.V., Fedorov I.R.** Blockchain technology in 5G networks. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2020, Vol. 20, No. 4, Pp. 472–484. DOI: 10.17586/2226-1494-2020-20-4-472-484

8. **Zheng Z., Xie S., Dai H., Chen X., Wang H.** An overview of blockchain technology: Architecture, consensus, and future trends. *2017 IEEE International Congress on Big Data* (*BigData Congress*), 2017, Pp. 557–564. DOI: 10.1109/BigDataCongress.2017.85

9. **Turginbayeva A.N., Tarabella A., Tanassoglo Y.** Use of blockchain technology in the highest education in Republic of Kazakhstan. *Bulletin of the Karaganda University*, 2018, Vol. 91, No. 3, Pp. 114–121.

10. **Kuo T.-T., Kim H.-E., Onho-Machado L.** Blockchain distributed ledger technologies for biomedical and health care applications. *Journal of the American Medical Informatics Association* (*JAMIA*), 2017, Vol. 24, No. 6, Pp. 1211–1220. DOI: 10.1093/jamia/ocx068

11. **Toilybayev A.Ye., Aitimov M.Zh., Bimuratkyzy Zh., Aitimova U.Zh.** Principles of the technology blockchain. *The Bulletin of Kazakh Academy of Transport and Communications named after M. Tynyshpayev*, 2017, Vol. 103, No. 4, Pp. 249–257.

12. **Sokolova T.N., Voloshin I.P., Petrunin I.A.** Pros and cons of the blockchain technology. *Ekonomicheskaia bezopasnost' i kachestvo* [*Economic Safety and Quality*], 2019, Vol. 34, No. 1, Pp. 49–52.

13. **Zharkova Yu.S.** Blockchain digital technologies: Advantages and disadvantages. *Zametki uchenogo* [*Scientist's Notes*], 2020, No. 10, Pp. 232–235.

14. **Dubovitskaya A., Xu Z., Ryu S., Schumacher M., Wang F.** Secure and trustable electronic medical records sharing using blockchain. *American Medical Informatics Association Annual Symposium Proceedings*, 2017, Pp. 650–659.

15. **Shafagh H., Burkhalter L., Hithnawi A., Duquennoy S.** Towards blockchain-based auditable storage and sharing of IoT data. *CCSW'17: Proceedings of the 2017 on Cloud Computing Security Workshop*, 2017, Pp. 45–50. DOI: 10.1145/3140649.3140656

16. **Ateniese G., Magri B., Venturi D., Andrade E.** Redactable blockchain – or – rewriting history in bitcoin and friends. *2017 IEEE European Symposium on Security and Privacy* (*EuroS&P*), 2017, Pp. 111–126. DOI: 10.1109/EuroSP.2017.37

17. **Kwon J., Buchman E.** Cosmos: A network of distributed ledgers, 2019. Available: https://resources.cryptocompare.com/asset-management/15/1662453504718.pdf (Accessed 10.05.2025).

18. **Kiayias A., Russell A., David B., Oliynykov R.** Ouroboros: A provably secure Proof-of-Stake blockchain protocol. *Advances in Cryptology – CRYPTO 2017*, 2017, Vol. 10401, Pp. 357–388. DOI: 10.1007/978-3-319-63688-7_12

19. **Reshi I.A., Sholla S.** IBF network: enhancing network privacy with IoT, blockchain, and fog computing on different consensus mechanisms. *Cluster Computing*, 2025, Vol. 28, Art. no. 208. DOI: 10.1007/s10586-024-05026-w

20. **Poon J., Buterin V.** Plasma: Scalable autonomous smart contracts, 2017. Available: https://plasma.io/plasma.pdf (Accessed 10.05.2025)

21. **Sharma A., Chauhan R., Gupta S., Kapruwan A.** Blockchain revolution in healthcare: A comprehensive survey. In: *Challenges in Information, Communication and Computing Technology* (eds. V. Sharmila, S. Kannadhasan, A. Rajiv Kannan, P. Sivakumar, V. Vennila), 2025, Pp. 218–223. DOI: 10.1201/9781003559085-38

22. **Maadallah Y., El Bouzekri El Idrissi Y., Baddi Y.** Enhancing IoT security through blockchain an in-depth analysis of the proof-of-work consensus mechanism. *EDPACS – Electronic Data Processing Audit, Control and Security*, 2025, Vol. 70, No. 5, Pp. 1–44. DOI: 10.1080/07366981.2025.2454095

23. **Ghorbian M., Ghobaei-Arani M.** Key concepts and principles of blockchain technology. *arXiv:2501.11707*, 2025. DOI: 10.48550/arXiv.2501.11707

24. **Bathini P.K., Manideep D.** Applications of artificial intelligence to blockchain consensus mechanisms: Using machine learning to make decentralized networks faster and more scalable. *2025 Global Conference in Emerging Technology* (*GINOTECH*), 2025, Pp. 1–6. DOI: 10.1109/GINOTECH63460.2025.11076838

25. **Soundararajan G., Tyagi A.K.** Blockchain technology: An introduction. In: *Blockchain Technology in the Automotive Industry* (eds. G. Yasin, A.K. Tyagi, T.A. Nguyen), 2025, Pp. 3–36. DOI: 10.1201/9781003450306

26. **Marwaha M., Bedi R.K., Gupta S.K.** An analysis of blockchain ecosystems: Understanding types, consensus models and security. *2025 12th International Conference on Computing for Sustainable Global Development* (*INDIACom*), 2025, Pp. 1–6. DOI: 10.23919/INDIACom66777.2025.11115579

27. **Niu G.** A Blockchain-based secure and privacy-preserving healthcare data management framework with SHA-256 and PoW consensus. *Informatica*, 2025, Vol. 49, No. 20, Pp. 149–162. DOI: 10.31449/inf.v49i20.8392

28. **Yu S., Qiao Y., Yang F., Bo J.** DPoW: A decentralized proof-of-work consensus mechanism for blockchain system. *Computer Networks*, 2025, Vol. 270, Art. no. 111490. DOI: 10.1016/j.comnet.2025.111490

INFORMATION ABOUT AUTHORS / СВЕДЕНИЯ ОБ АВТОРАХ

Daniil D. Razuvaev
Разуваев Даниил Дмитриевич
E-mail: Razuvaev_DD@mail.ru

Sergey M. Ustinov
Устинов Сергей Михайлович
E-mail: usm50@yandex.ru
ORCID: https://orcid.org/0000-0003-4088-4798

Submitted: 30.06.2025; Approved: 25.11.2025; Accepted: 19.12.2025.
Поступила: 30.06.2025; Одобрена: 25.11.2025; Принята: 19.12.2025.

Computing, Telecommunication and Control, 2025, Vol. 18, No. 4, Pp. 102–111.
Информатика, телекоммуникации и управление. 2025. Том 18, № 4. С. 102–111.

Simulations of Computer, Telecommunications and Control Systems

Моделирование вычислительных, телекоммуникационных и управляющих систем

Research article
DOI: https://doi.org/10.18721/JCSTCS.18409
UDC 004.496, 616.8-085.851

APPLICATION OF VR-TECHNOLOGIES FOR NEUROREHABILITATION OF PATIENTS WITH MOTOR AND COGNITIVE FUNCTION DISORDERS

N.N. Reshetnikova[1] ✉, *A.S. Kuzmin*[2], *A.V. Nikitin*[1], *I.S. Karamyshev*[1]

[1] Saint-Petersburg State University of Aerospace Instrumentation, St. Petersburg, Russian Federation;

[2] "V.A. Almazov National Medical Research Center" of the Ministry of Health of the Russian Federation, St. Petersburg, Russian Federation

✉ reni_07@list.ru

Abstract. The article explores the use of developed interactive 3D application with VR technology support, designed for the rehabilitation of patients after neurological disorders that cause motor and cognitive impairments. The neurorehabilitation hardware and software system is implemented on the Unity platform and includes an interactive 3D application and a Quest 3 VR headset. The VR headset consists of a virtual reality helmet and two controllers for the right and left hands, respectively, and supports highly accurate controller tracking using built-in cameras equipped with LiDAR technology. The causes of unilateral spatial neglect syndrome, or hemispatial neglect, and the possibility of its rehabilitation are considered. The hardware and software system developed by the authors enhances patient motivation through gamification and adaptive scenarios, allowing them to select and visualize various 3D scenes. When immersed in a realistic virtual environment, the patient acts instinctively. This helps develop hand motor skills, a sense of balance and spatial navigation abilities. Furthermore, the hardware and software system allows for the convenient storage of statistical data on the progress of prescribed procedures and can be used for further analysis during rehabilitation process.

Keywords: cognitive functions, motor functions, hemispatial neglect , rehabilitation, virtual reality, interactive 3D graphics, game engine

Citation: Reshetnikova N.N., Kuzmin A.S., Nikitin A.V., Karamyshev I.S. Application of VR-technologies for neurorehabilitation of patients with motor and cognitive function disorders. Computing, Telecommunications and Control, 2025, Vol. 18, No. 4, Pp. 102–111. DOI: 10.18721/JCSTCS.18409

Научная статья
DOI: https://doi.org/10.18721/JCSTCS.18409
УДК 004.496, 616.8-085.851

ПРИМЕНЕНИЕ VR-ТЕХНОЛОГИЙ ДЛЯ НЕЙРОРЕАБИЛИТАЦИИ ПАЦИЕНТОВ С НАРУШЕНИЕМ ДВИГАТЕЛЬНЫХ И КОГНИТИВНЫХ ФУНКЦИЙ

Н.Н. Решетникова[1] ✉ , *А.С. Кузьмин*[2], *А.В. Никитин*[1] , *И.С. Карамышев*[1]

[1] Санкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, Российская Федерация;

[2] «Национальный медицинский исследовательский центр имени В.А. Алмазова» Министерства здравоохранения Российской Федерации, Санкт-Петербург, Российская Федерация

✉ reni_07@list.ru

Аннотация. В статье рассматривается применение разработанного интерактивного 3D-приложения с поддержкой VR-технологий, направленного на реабилитацию пациентов после неврологических заболеваний, вызывающих нарушения двигательных и когнитивных функций. Программно-аппаратная система нейрореабилитации реализована на платформе Unity и включает в себя интерактивное 3D-приложение и VR-гарнитуру – Quest 3. VR-гарнитура состоит из шлема виртуальной реальности и двух контроллеров для правой и левой руки соответственно, поддерживается достаточно точный трекинг контроллеров с помощью встроенных камер, оснащенных технологией LiDAR. Рассмотрены причины развития одностороннего пространственного игнорирования, или синдрома неглекта, и возможность его реабилитации. Разработанный авторами программно-аппаратный комплекс повышает мотивацию пациентов за счет геймификации и адаптивных сценариев, позволяя выбирать и визуализировать различные 3D-сцены. При погружении в виртуальную среду, приближенную к реальности, пациент действует инстинктивно. Таким образом развивается моторика рук, чувство равновесия, навыки перемещения в пространстве. Кроме того, программно-аппаратный комплекс позволяет сохранять в удобном цифровом виде статистические данные о ходе выполнения назначенных процедур и может использоваться для дальнейшего анализа при прохождении реабилитации.

Ключевые слова: когнитивные функции, двигательные функции, синдром неглекта, реабилитация, виртуальная реальность, интерактивная 3D-графика, игровой движок

Для цитирования: Reshetnikova N.N., Kuzmin A.S., Nikitin A.V., Karamyshev I.S. Application of VR-technologies for neurorehabilitation of patients with motor and cognitive function disorders // Computing, Telecommunications and Control. 2025. Т. 18, № 4. С. 102–111. DOI: 10.18721/JCSTCS.18409

Introduction

Today, in rehabilitation medicine for neurologists, instructors in adaptive physical education, addressing motor and cognitive impairments in patients who have suffered a stroke is of particular importance. In recent years, rehabilitation of patients after a stroke, has evolved significantly, reaching a fundamentally new level. Traditional methods (manual therapy, exercise equipment, physiotherapy, medications) have been supplemented by advanced developments based on digital technologies [1].

The article examines in detail the approach to the design and implementation of a hardware and software neurorehabilitation system based on the Unity game engine, including an interactive 3D application and a Quest 3 VR headset. The VR headset consists of a VR helmet and two controllers,

for the right and left hands, respectively, with fairly accurate controller tracking enabled by built-in cameras equipped with LiDAR technology.

In particular, the article provides a detailed description and analysis of the causes of the unilateral spatial neglect syndrome, or hemispatial neglect, and proposes a method for its diagnosis and rehabilitation using VR technologies. The hardware and software complex developed by the authors increases patient motivation through gamification and adaptive scenarios, enabling users to select and visualize various 3D scenes. When immersed in a realistic virtual environment, patients act instinctively. This promotes the development of hand motor skills, a sense of balance and spatial navigation abilities. In addition, the hardware and software complex allows for highly accurate collection of statistical data on the progress of prescribed procedures, storing them in a convenient digital form and using them for further analysis during rehabilitation.

Hemispatial neglect

Strokes lead to various disorders of motor and cognitive functions in the human body, as well as changes in emotional regulation. One consequence of brain damage is the development of hemispatial neglect [2, 3]. The term "neglect" derives from Latin *neglēctus*, perfect passive participle of *neglegō*, which means "disregard". It is a neuropsychological condition in which a deficit in attention and awareness towards the side of space opposite brain damageis observed. Fig. 1 shows manifestations of hemispatial neglect when the right or left hemisphere of the brain is damaged.

According to statistics, the overall prevalence of hemispatial neglect is up to 82% in post-stroke patients, while an average of 50% of patients still have manifestations of neglect. Visual neglect manifests as the patient's inability to detect objects located in the contralateral space. The patient acts as if unable to perceive anything on the left or right side, although in fact, vision is intact. Auditory neglect manifests as the patient's inability to perceive sounds in one ear, even when the information is important. Tactile neglect manifests as the patient's inability to respond to touch and other stimuli on one side of the body.

It should be noted that the wide variation in reported statistics is explained by the differences in approaches to understanding neglect and its underlying brain mechanisms, as well as the complex structure of the syndrome itself and the lack of specialized psychodiagnostic tools.

As a polymodal syndrome with a complex structure, hemispatial neglect presents itself with highly diverse manifestations. A patient may exhibit one, several or all signs of neglect. Among the clinical features of neglect, the following can be distinguished:

- hemineglect (failure to respond appropriately to surrounding stimuli such as approaching people or various sounds);

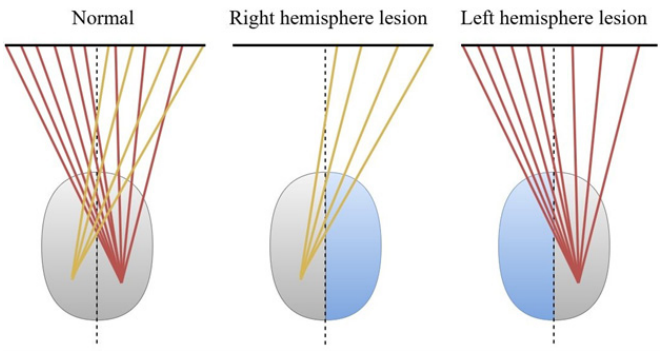


Fig. 1. Manifestations of hemispatial for the right or left hemisphere of the brain

• tactile neglect (failure to respond appropriately to tactile stimuli when both sides of the body are touched simultaneously);

• visual neglect (failure to respond appropriately to visual stimuli when both visual fields are stimulated simultaneously);

• alloesthesia (a condition in which a patient responds to stimuli presented to one side of their body as if they were presented at the opposite side);

• anosognosia (a condition in which a patient with a disability is cognitively unaware of having it);

• asomatognosia (failure to recognize or be aware of one's part of the body).

European Handbook of Neurological Management provides an overview of the methods and techniques for hemispatial neglect rehabilitation. There is evidence for the effectiveness of multiple approaches in reducing hemispatial neglect manifestations:

• combined training of visual scanning, reading, copying and figure description;
• spatiomotor or visuo-spatiomotor cueing;
• visual cueing with kinetic stimuli;
• video and visuomotor feedback;
• training of sustained attention, increasing of alertness or cueing of spatial attention;
• influencing multisensory representations;
• prism goggles;
• forced use of the left visual hemifield or left eye;
• computer training [4–6].

In addition, the use of biofeedback, such as eye-tracking [7] or motion capture [8], enables real-time adaptation of the rehabilitation program based on the patient's individual performance.

Recently, computer-based cognitive training has gained considerable popularity. To date, a substantial body of evidence has been accumulated, leading specialists to conclude that computer-based cognitive training is effective for patients with focal brain lesions[1].

Upon closer examination, it is evident that computer games actively engage higher mental functions such as attention, executive control, task maintenance, differentiation and memory. Therefore, the gamification of rehabilitation procedures, that allows for simulating everyday or extreme situations, can enhance patient motivation [9].

Two internationally standardized techniques widely accepted for correcting hemispatial neglect are combined training of visual scanning, reading, copying and figure description and spatiomotor or visuo-spatiomotor cueing. These methods aim to improving spatial perception, orientation and ultimately the functional independence of patients with hemispatial neglect[2].

The literature to date describes various methods of restorative neuropsychological intervention for patients with hemispatial neglect. In addition to developing specialized computer-based cognitive trainings for patients with cognitive impairments, existing computer games can also be used. Most of these techniques focus on rehabilitating sensory and motor functions [10].

For patients with hemispatial neglect, the rehabilitation involves modifying the physical environment and engaging the patient in virtual environments through alternating hand use. The key requirement is the active involvement of the affected side of the body in purposeful activities.

Rehabilitation 3D-system with VR-technology support

Considering that the purpose of the application is the rehabilitation of patients with complete or partial impairment of cognitive and motor functions, the rehabilitation process must be carried out under

[1] Home – Хабилект, Available: https://habilect.com/en/home_en/ (Accessed 15.01.2026); NIRVANA | Sensory and interactive room | BTS Bioengineering, Available: https://www.btsbioengineering.com/products/nirvana/ (Accessed 15.01.2026); VRRS Evo – Khymeia, Available: https://khymeia.com/en/products/vrrs-evo/ (Accessed 15.01.2026); Rehabilitation, Available: https://www.who.int/news-room/fact-sheets/detail/rehabilitation (Accessed 15.01.2026)

[2] Home – Хабилект, Available: https://habilect.com/en/home_en/ (Accessed 15.01.2026)

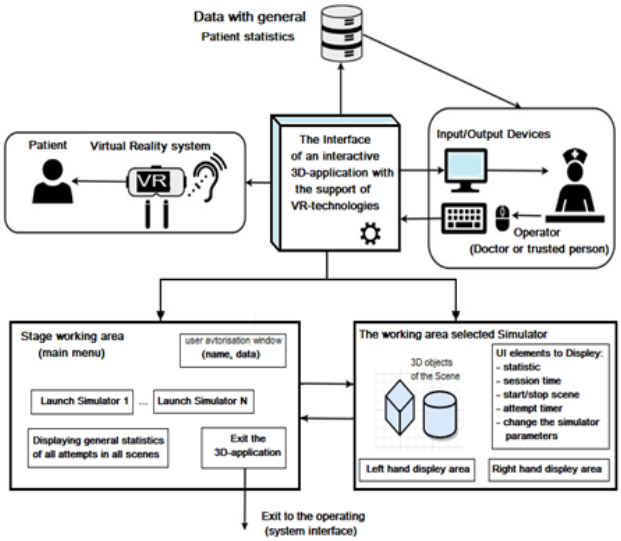


Fig. 2. General diagram of the diagnostic and rehabilitation system

the supervision of an operator. The operator monitors the progress of the procedure and adjust the parameters in accordance with the patient's limitations.

The operator can be the attending physician or another trained person who can oversee the session and adhere to all necessary rehabilitation guidelines. The operator is responsible for user authorization, selecting the type of simulator, adjusting 3D scene parameters, saving statistics and exporting the results to a separate file for further analysis if necessary.

The operator controls the procedure from a desktop computer or laptop, using a mouse and keyboard(or a touchpad in case of a laptop).

The procedure is displayed on the operator's monitor in full synchronization with the virtual environment seen by the patient. It also provides an interface for interacting with the application.

The VR helmet displays only the 3D image of the environment, with which the patient interacts using hand-held controllers (in one or both hands). The helmet also includes built-in speakers for audio playback.

A general diagram of the diagnostic and rehabilitation system is shown in Fig. 2.

The main functions required in the application are:

- authorization of the patient by indicating their full name;
- simulator selection, including one of the 3D scenes from the list, with the scenario appropriate for rehabilitation;
- ability to start/stop the simulator session, restricted to the operator;
- tracking of time spent in the VR system, as well as calculation of statistics for each session;
- ability to export statistics from the 3D scene for deeper analysis;
- ability to restart the simulator, resetting the timer and attempt statistics;
- ability for the operator to adjust exercise parameters for specific simulators.

The Quest 3 modern VR system was chosen as the hardware component of the project. The system includes a VR headset and two controllers (for the right and left hand, respectively) (Fig. 3).

The patient benefits from high device quality, comfortable ergonomics, a stable connection to a desktop computer or laptop, high image clarity with adjustable focal length, and fairly accurate controller tracking using built-in cameras equipped with LiDAR technology (for digitizing the surrounding space).

The main characteristics of the VR headset are presented in Table 1.

Fig. 3. Quest 3 VR System

Table 1

Main characteristics of the VR system

Characteristic	Value
Per-eye resolution	2064×2208 pixels
Memory	8 GB
Storage	128 GB
Display refresh rate	90–120 Hz
Field of view	104°H/96°V
Connectivity	Bluetooth 5.2, Wi-Fi 6E

The minimum system requirements are presented in Table 2.

Table 2

Minimum system requirements

OS	Windows 10, Windows 11
CPU	AMD Ryzen 5 1500, Intel Core i5-4590
Video card	AMD R9 290, NVIDIA GTX 970
RAM	8 GB

To facilitate further testing, the headset is additionally equipped with a special accessory that ensures secure fixation on the patient's head when put on.

Since hand movements of patients undergoing post-illness rehabilitation are not always predictable, the use of special controller straps is essential.

Scenarios

Developing scenarios for patient interaction with the virtual environment is one of the most important stages of designing an interactive 3D application with VR technology support aimed at rehabilitation and restoration of lost cognitive and motor skills. At this stage, it is necessary to define the logic of how the patient will interact with the events in each specific scene, what actions they will perform, and how the system should respond.

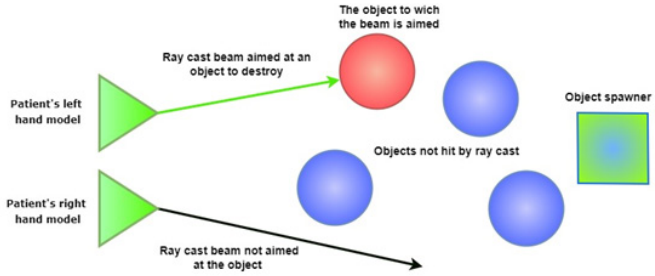


Fig. 4. Simplified diagram of patient's interaction with the virtual environment

Consider a rehabilitation scenario using a specific example of one of the simulators implemented in the diagnostic and rehabilitation system.

After the operator (the attending physician) selects a simulator's 3D scene from the main menu, the image in the VR helmet changes and the patient enters the simulated virtual space. The patient surveys the environment, verifies the representation of their right- and left-hand models in space, and then confirms readiness to begin interaction with the simulator. Once readiness is confirmed, the operator starts the simulator mechanics using the available interface.

A timer built into the scene starts counting down, and within a defined area of the environment, objects begin to appear at a given interval, moving toward the patient.

In turn, the patient, using beams emitted from their hand models, must aim at the approaching objects. Accurate targeting changes the object's texture. By pressing a certain trigger on the controller, the patient can destroy the approaching object. Each successful destruction adds one point to the statistics counter. Upon reaching a certain result or a time limit, the patient ends the session, and the operator stops the simulator.

The operator logs the session statistics and, if necessary, resets the indicators.

A simplified diagram of the patient's interaction with the surrounding virtual environment is shown in Fig. 4.

Implemented system on Unity platform

The diagnostic and rehabilitation system is implemented as an interactive 3D application with VR technology support on the Unity platform[3].

The following software tools were used:

- Unity 6.0 (6000.0.37f1) – a modern game engine designed for developing interactive applications, games and visualizations, used as the main tool for developing mechanics, scene environments and interfaces.
- SteamVR Plugin 2.8.0 (SDK 2.0.10) – a key tool for VR development in Unity, used as an integrated tool in the project.
- Visual Studio 2022 – a modern integrated development environment for creating software solutions in a variety of languages, including C# for scripting.
- Excel 2021 – a spreadsheet application designed for collecting, analyzing and processing data in various formats, used to collect and analyze the resulting statistical data.

The final appearance of the operator interface for the simulator, implemented in the Unity project, the scenario for which is given above, is shown in Fig. 5.

[3] Платформа Unity для разработки в реальном времени | Движок для 3D, 2D, VR и AR, Available: https://unity.com/ru (Accessed 15.01.2026); GitHub – gkngkc/UnityStandaloneFileBrowser: A native file browser for unity standalone platforms, Available: https://github.com/gkngkc/UnityStandaloneFileBrowser (Accessed 15.01.2026)

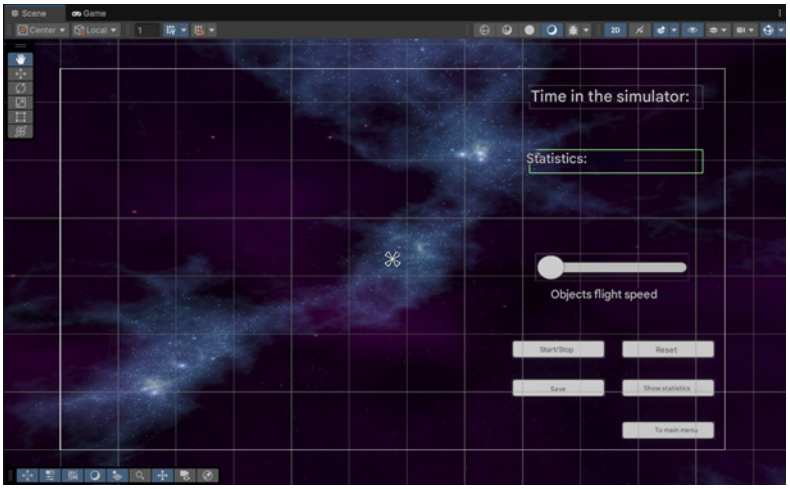


Fig. 5. Operator interface for the simulator

The following interface elements have been added and configured:

- Time spent in the session.
- Statistics of destroyed objects.
- Slider for changing the speed of the objects' movement.
- "Start/Stop" button to start or pause the movement of objects, timer or statistics.
- "Reset" button to reset the timer and statistics.
- "Save" button to save the results.
- "Show Statistics" button to show general statistics.
- "To Main Menu" button to get the main menu.

For general statistics, the following interface elements have been added and configured:

- Statistics of each simulator scene.
- "Clear Statistics" button to clear the general statistics.
- "Download to Excel" button to export statistics to a CSV file.
- "To Main Menu" button to get the main menu.

After the interactive 3D application was created, tests were run to verify its functionality. Testing was conducted in accordance with the intended application logic and defined objectives:

- Capturing the view from the VR headset camera and transmitting the image to the operator's display.
- Correct patient interaction with the environment using controllers.
- Performance of the operator interface.
- Interaction between scenes.
- Exporting statistics data to a CSV file.

After the simulator scene loads, the operator's screen displays the image from the patient's headset camera, along with control elements.

Pressing the "Start/Stop" button launches the simulator. Objects appear, which the patient must destroy with a specially directed green beam. A countdown timer starts, and statistics on the number of objects destroyed are recorded.

The result of the functionality described above is shown in Fig. 6.

Clicking the "Save" button saves the obtained data. Adjusting the 3D scene parameters increases the object speed using the corresponding slider.

If a new session is launched and its result is saved, the "Show Statistics" button can be slicked to show the general statistics.

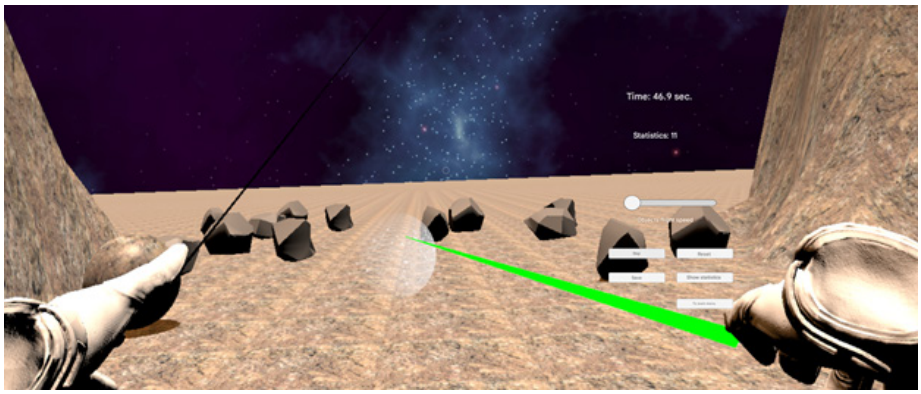


Fig. 6. Mirroring patient's display (Quest 3) to the operator's monitor

To export data from the general statistics to a separate CSV file the “Export to Excel” button should be clicked.

All functions implemented in the application were verified during testing.

The final build of the application takes up 271 MB of disk space.

Frame rate of up to 144 FPS at a resolution of 2064x2208.

Response time to patient actions is 0.1 ms.

Application optimization ensures moderate load on the hardware of a PC or laptop.

Conclusion

In conclusion, the following points should be noted:

1. Experimental testing of an interactive 3D application with VR support for the rehabilitation of patients with motor impairments demonstrates that VR technologies not only enhance patient immersion in the rehabilitation process, but also enable an individualized approach to correcting motor and cognitive functions based on medical indicators.

2. Impaired visual perception affects body function and structure, which in turn limits activity and participation in daily life. Visual impairment has significant physical, social, emotional, behavioral and economic consequences. It can have a pronounced impact on patients’ disability one year after a stroke, making the timely use of VR technologies in neurorehabilitation a relevant direction of work.

3. The core objective of neurorehabilitation is the restoration of cognitive and sensorimotor functions essential for daily living and their impairment significantly affects overall quality of life. Therefore, advanced technologies such as VR and augmented reality open up new possibilities for creating immersive environments where patients can perform exercises as simple and understandable tasks in a safe and controlled environment, fostering a sense of “presence” in the virtual space.

4. Equally important is to record patient’s progress and, based on the data obtained, to understand whether the chosen rehabilitation scenario is appropriate or whether adjustments to the sequence of procedures are required.

REFERENCES

1. **Kuznetsov A.A., Petrenko T.S.** Rehabilitation of cognitive functions in a virtual environment: A systematic review. *Ural Medical Journal*, 2024, Vol. 23, No. 6, Pp. 91–107. DOI: 10.52420/umj.23.6.91

2. **Hankey G.J.** European Handbook of Neurological Management. January 2008. *Practical Neurology*. 8(1): 69–69. DOI:10.1136/jnnp.2007.139709. Available: https://www.researchgate.net/publication/274531987 (Accessed 18.01.2026)

3. Hankey's Clinical Neurology. Third edition. Edited by: Gorelick P.B., Testai F.D., Hankey G.J., Wardlaw J.M. / Taylor & Francis Group, 2021. 953p.

4. *European handbook of neurological management* (eds. N.E. Gilhus, M.P. Barnes, M. Brainin), 2nd ed. Chichester (UK): John Wiley & Sons, 2011, Available: http://ndl.ethernet.edu.et/bitstream/123456789/42514/1/319.pdf (Accessed 16.01.2026).

5. **Nikitayeva E.V.** *Neyropsikhologicheskaia reabilitatsiia patsientov s sindromom neglekta* (*sindromom zritel'no-prostranstvennogo ignorirovaniia*) [*Neuropsychological rehabilitation of patients with neglect's syndrome* (*unilateral visuospatial neglect syndrome*)]. Kazan: Buk, 2021.

6. **Akhutina T., Varako N.A., Grigor'eva V., Mikadze Iu.V., Skvortsov A.A., Fufaeva E.V.** *Kliniko-psikhologicheskaia diagnostika i kliniko-psikhologicheskaia reabilitatsiia patsientov s narusheniiami reguliatornykh funktsii pri povrezhdeniiakh golovnogo mozga* [*Kliniko-psikhologicheskaia diagnostika i kliniko-psikhologicheskaia reabilitatsiia patsientov s narusheniiami reguliatornykh funktsii pri povrezhdeniiakh golovnogo mozga*]. Moscow: Soiuz reabilitologov Rossii, 2014.

7. **Ruiz-Beltrán C.A., Romero-Garcés A., González M., Pedraza A.S., Rodríguez-Fernández J.A., Bandera A.** Real-time embedded eye detection system. *Expert Systems of Application*, 2022, Vol. 194, Art. no. 116505. DOI: 10.1016/j.eswa.2022.116505

8. **Roberts E.** Are there any free motion capture software in 2025. Available: https://remocapp.com/blog/posts/1514/are-there-any-free-motion-capture-software-in-2025 (Accessed 15.01.2026)

9. **Buchackij K.V.** *Nejrovizual'naya diagnostika i nejroreabilitaciya bol'nyh gemispacial'nym neglektom s pomoshch'yu innovacionnyh komp'yuternyh tekhnologij* [*Izbrannye voprosy nejroreabilitacii: materialy VII mezhdunarodnogo kongressa «Nejroreabilitaciya — 2015»*]. Moscow: Soiuz reabilitologov Rossii, 2015.

10. **Nikitin A.V., Reshetnikova N.N., Main E.E., Sitnikov I.A., Ageev A.S., Deviatov A.A., Kataranov M.V., Kukhar' M.A., Nikulin O.V.** *Metavselennaia: Razrabotka i ispol'zovanie na primere GUAP* [*Metaverse: Development and Use at GUAP*]. St. Petersburg: GUAP, 2024.

INFORMATION ABOUT AUTHORS / СВЕДЕНИЯ ОБ АВТОРАХ

Nina N. Reshetnikova
Решетникова Нина Николаевна
E-mail: reni_07@list.ru
ORCID: https://orcid.org/0009-0002-4444-1249

Alexey S. Kuzmin
Кузьмин Алексей Сергеевич
E-mail: kuzmin_as@almazovcentre.ru

Alexander V. Nikitin
Никитин Александр Васильевич
E-mail: nike51@mail.ru
ORCID: https://orcid.org/0000-0002-8086-2633

Igor S. Karamyshev
Карамышев Игорь Сергеевич
E-mail: karamyshevigor96@gmail.com

Submitted: 12.10.2025; Approved: 30.11.2025; Accepted: 08.12.2025.
Поступила: 12.10.2025; Одобрена: 30.11.2025; Принята: 08.12.2025.

Computing, Telecommunication and Control, 2025, Vol. 18, No. 4, Pp. 112–122.
Информатика, телекоммуникации и управление. 2025. Том 18, № 4. С. 112–122.

System Analysis and Control
Системный анализ и управление

Research article
DOI: https://doi.org/10.18721/JCSTCS.18410
UDC 519.8

MULTI-CRITERIA CONTROL OF LARGE-SCALE NONLINEAR DYNAMICAL SYSTEMS WITHOUT LINEARIZATION, BASED ON LYAPUNOV FUNCTIONS

M.I. Fershtadt ✉ , *V.N. Shashikhin*

St. Petersburg State Polytechnical University,
St. Petersburg, Russian Federation

✉ tral1930@mail.ru

Abstract. This paper proposes a numerical control method for large-scale nonlinear dynamical systems, focused on maintaining stability without using linearization. The approach under study is based on the principles of multi-criteria optimization, where the stability of the system is directly included in the vector of target criteria through Lyapunov functions. This allows us not only to minimize deviations from the target states and energy consumption for control, but also to guarantee the asymptotic stability of the system under arbitrary initial conditions. A mathematical formulation of the problem is presented, a discrete numerical control scheme is developed, and a scalarization strategy is proposed that provides an approximation to Pareto-optimal solutions. A series of numerical experiments implemented in Python has been conducted, confirming the effectiveness of the method using examples of both single- and multi-agent systems. The results demonstrate the stable behavior of the trajectories, a decrease in the Lyapunov function over time and correct operation even with strong nonlinearity of the model.

Keywords: nonlinear systems, stability, Lyapunov function, multi-criteria optimization, distributed control, numerical simulation, control without linearization

Citation: Fershtadt M.I., Shashikhin V.N. Multi-criteria control of large-scale nonlinear dynamical systems without linearization, based on Lyapunov functions. Computing, Telecommunications and Control, 2025, Vol. 18, No. 4, Pp. 112–122. DOI: 10.18721/JCSTCS.18410

Научная статья
DOI: https://doi.org/10.18721/JCSTCS.18410
УДК 519.8

МНОГОКРИТЕРИАЛЬНОЕ УПРАВЛЕНИЕ КРУПНОМАСШТАБНЫМИ НЕЛИНЕЙНЫМИ ДИНАМИЧЕСКИМИ СИСТЕМАМИ БЕЗ ЛИНЕАРИЗАЦИИ НА ОСНОВЕ ФУНКЦИЙ ЛЯПУНОВА

М.И. Ферштадт ✉ , *В.Н. Шашихин*

Санкт Петербургский государственный политехнический университет,
Санкт-Петербург, Российская Федерация

✉ tral1930@mail.ru

Аннотация. В данной статье предлагается метод численного управления крупномасштабными нелинейными динамическими системами, ориентированный на сохранение устойчивости без использования линеаризации. Исследуемый подход опирается на принципы многокритериальной оптимизации, где устойчивость системы напрямую включается в вектор целевых критериев посредством функций Ляпунова. Это позволяет не только минимизировать отклонения от целевых состояний и энергозатраты на управление, но и гарантировать асимптотическую устойчивость системы при произвольных начальных условиях. Представлена математическая постановка задачи, разработана дискретная численная схема управления и предложена стратегия скаляризации, обеспечивающая приближение к Парето-оптимальным решениям. Проведена серия численных экспериментов при помощи языка программирования Python, результаты моделирования представлены в виде графиков, подтверждающих эффективность метода на примере как одиночной, так и мультиагентной системы. Результаты демонстрируют устойчивое поведение траекторий, уменьшение функции Ляпунова во времени и корректную работу даже при сильной нелинейности модели.

Ключевые слова: нелинейные системы, устойчивость, функция Ляпунова, многокритериальная оптимизация, распределенное управление, численное моделирование, управление без линеаризации

Для цитирования: Fershtadt M.I., Shashikhin V.N. Multi-criteria control of large-scale nonlinear dynamical systems without linearization, based on Lyapunov functions // Computing, Telecommunications and Control. 2025. Т. 18, № 4. С. 112–122. DOI: 10.18721/JCSTCS.18410

Introduction

Modern technical, cyber-physical and social systems are characterized by high dimensionality, a complex structure of interactions between components and pronounced nonlinear dynamics [1]. Such systems include intelligent power grids, multi-agent robotic platforms, traffic management systems, distributed computing complexes and others [2, 3]. Increasing complexity creates the need for new control approaches, the necessity to simultaneously consider multiple quality criteria, interactions between subsystems, and stability requirements [4, 5].

Traditional control methods based on linearization of the model demonstrate limited applicability under conditions of rapid transients, saturation of control actions and variable system structure. In particular, linearization can significantly distort the real properties of the system, reducing control accuracy and depriving the developer of stability guarantees outside a small neighborhood of the equilibrium [6].

In these conditions, multi-criteria optimization (MCO) without linearization becomes particularly relevant, allowing for simultaneous consideration of dynamic and energy characteristics, as well as stability characteristics [7–10]. However, most existing approaches to MCO treat stability as an additional

constraint or a posteriori verification of results, without including it in the optimization structure itself. This creates a gap between the formal optimization task and the actual requirements for the safety and stability of the system [11–14].

The present work introduces an approach in which the Lyapunov function is included directly in the vector of criteria for MCO. This allows for integrating stability requirements into the optimization process. This approach can be applied in both centralized and distributed control systems.

Mathematical model and formalization of multi-criteria control

Structure and properties of the controlled system

By a large-scale nonlinear system, we mean a set of M interacting subsystems $\sum i$, each of which has its own dynamics and interacts with others through a limited set of connections. The connectivity of the system is represented by an oriented graph $G = (V, E)$, where the vertices $V = \{1, 2, ..., M\}$ correspond to subsystems, and the edges $E \subset V \times V$ determine the structure of the interaction [15, 16].

The dynamics of each subsystem $\sum i$ is given by a system of differential equations:

$$\dot{x}_i = f_i\left(x_i,\ u_i,\ x_{N_i}\right),\quad y_i = h_i\left(x_i\right),$$

where: $x_i \in \mathbb{R}_i^n$ is the state vector; $u_i \in \mathbb{R}_i^m$ is the control input; $N_i \subset V$ is a set of neighboring agents with which there is a connection that affects subsystems $\sum i$; x_{N_i} is the vector of all neighboring states; f_i, h_i is the continuous functions, and in some cases, non-linear ones.

The state of the entire system:

$$x = \left[x_1^T,\ x_2^T,\ \ldots,\ x_M^T\right]^T \in \mathbb{R}^n,\quad n = \sum_{i=1}^{M} n_i.$$

Problem formulation for multi-criteria control

Modern control tasks require simultaneous satisfaction of several, often contradictory, goals: minimizing deviations from the trajectory, reducing control effort and ensuring stability and coordination between system components [17–20]. Formally, such a task is formulated as a MCO with a vector objective function:

$$J\left(x,\ u\right) = \begin{bmatrix} J_1\left(x,\ u\right) \\ J_2\left(x,\ u\right) \\ \vdots \\ J_K\left(x,\ u\right) \end{bmatrix},\quad J_K : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}.$$

The desired control $u(t)$ must be Pareto-optimal for the vector of criteria $J(x(t))$, $u(t))$ under the following constraints:

1. Equations of motion of each subsystem:

$$\dot{x}_i = f_i\left(x_i,\ u_i,\ x_{N_i}\right).$$

2. Initial conditions:

$$x_i\left(0\right) = x_{i0}.$$

3. Functional constraints on control and state:

$$u_i(t) \in U_i, \quad x_i(t) \in X_i, \quad \forall t \in [0, T].$$

Stability as an integrated criterion: the Lyapunov function

The key feature of the proposed approach is the direct inclusion of the stability requirement in the objective function of the optimization problem. This is achieved through a Lyapunov function $V(x)$ that satisfies the classical conditions:

$$V(x) > 0 \text{ at } x \neq 0, \quad V(0) = 0;$$
$$\dot{V}(x) = \nabla V(x)^T f(x, u) < 0.$$

Using the Lyapunov function as one of the optimization criteria allows one not only to ensure the stability of the system, but also to integrate the quality of the transition process into the objective vector. The structure of the objective functions may include:

- control accuracy:

$$J_1 = \int_0^T \left\| x(t) - x_{ref}(t) \right\|^2 dt,$$

where $x_{ref}(t)$ is the desired trajectory;

- control effort:

$$J_2 = \int_0^T \left\| u(t) \right\|^2 dt.$$

- stability:

$$J_3 = V(x(T)).$$

- convergence rate:

$$J_4 = \int_0^T -V(x(t))\, dt.$$

For systems with polynomial nonlinearity, it is possible to use sums of squares of polynomials, which makes it possible to efficiently construct and verify Lyapunov functions via semidefinite programming:

$$V(x) \in \sum_{[x]}, \quad -\dot{V}(x) = -\nabla V(x)^T f(x) \in \sum_{[x]}.$$

Distributed problem formulation

When implementing control in a distributed environment, each agent solves a local optimization problem, taking into account its objectives and coordination with its neighbors. In this case, the optimization takes the following form:

$$J_i = \sum_{k=0}^{N-1} \left[V_i(x_i^k) + \rho \left\| x_i^k - \bar{x}_{N_i}^k \right\|^2 + \beta \left\| u_i^k \right\|^2 \right],$$

where: $\bar{x}_{N_i}^k$ is the weighted average of the states of the neighbors; ρ, β are the coefficients of coordination and control effort.

Information exchange between agents takes place according to a certain topology of the graph $G = (V, E)$, and coordination can be implemented either via consensus iterations or via the Alternating Direction Method of Multipliers (ADMM).

Development of a numerical method for multi-criteria optimization

Requirements for the numerical algorithm

The following key requirements apply for the practical implementation of the proposed optimization method:

1) working with the original nonlinear model – without prior linearization;
2) multi-criteria capability – the presence of several goals in the objective function, including stability;
3) numerical stability – maintaining the operability of the method under coarse approximations;
4) distributed implementation – the ability to perform optimization independently for each subsystem with coordination;
5) flexibility – the ability to adapt to the constraints and changing structure of the model.

These requirements impose restrictions on the choice of integration method, convergence criteria and the structure of the computational process.

Discretization of dynamics

For numerical analysis, the continuous dynamics of the system is discretized in time. The simplest Euler scheme is used:

$$x_{k+1} = x_k + \Delta t \times f\left(x_k,\, u_k\right),\quad k = 0,\, \ldots,\, N-1,$$

where: Δt is the time step; N is the number of discretization steps.

With this scheme, the state sequence x_k is generated sequentially using the control values u_k.

In the future, the Euler scheme can be replaced by more stable schemes – Runge–Kutta or the backward Euler scheme.

The system is constrained by physical bounds, the states and controls are clipped to prescribed ranges:

$$x_k \in X = \left[-X_{\max},\, X_{\max}\right],\quad u_k \in U = \left[-u_{\max},\, u_{\max}\right].$$

Structure of the scalar functional

For numerical solution, the multi-criteria problem is scalarized – one combined functional is formed [21, 22]:

$$J\left(u\right) = \alpha \sum_{k=0}^{N-1} V\left(x_k\right) \times \Delta t + \beta \sum_{k=0}^{N-1} u_k^2 \times \Delta t,$$

where: $V(x_k)$ is the the Lyapunov function at step k; u_k is the control input.

The parameters α, β are set by the user and reflect the priority between stability and control effort.

This approach has the following advantages:

1. **Stability is included in the objective function**, which guarantees a decrease in $V(x)$ along the optimal trajectory. It is important to note here that stability is treated not as a hard **constraint**, but as one of the criteria.

2. After discretization and scalarization, the usual function $J(u)$ is obtained and **the problem can be solved numerically using Python** and scipy.optimize.minimize or other standard nonlinear optimizers.

3. **The algorithm is simple to implement in real time**. If N is chosen to be small, control can be computed in a few seconds, without complex squares of polynomials or semidefinite programming solvers, in the proposed algorithm the control action is updated at each control cycle.

Integration of Lyapunov conditions

To integrate stability into a numerical algorithm, it is necessary:

- fix the parameter $\delta > 0$;
- introduce a restriction:

$$\dot{V}(x_k) = \nabla V(x_k)^T f(x_k, u_k) \le -\delta \|x_k\|^2.$$

If $V(x)$ is a quadratic or polynomial function (for example, $V(x) = x^T P x$), then verification of this condition is possible numerically. At the same time, the feasibility of the optimization problem remains, even if the exact minimum is not achieved, due to the preservation of the stability property.

Optimization algorithm

The general numerical scheme is as follows:

1. Initialization:
- the initial state x_0 is set;
- the parameters N, Δt, α, β are set;
- an approximation of $f_d(x, u)$ is chosen (for example, using the Euler method);
- an initial control guess $u^{(0)}$ is selected (for example, zeros).
2. Optimization:
- the objective functional $J(u)$ is constructed;
- the optimization method is chosen (gradient, quasi-Newton, evolutionary);
- if necessary, restrictions and filtering of values are introduced.
3. Stability check:
- it is necessary to make sure that the values of $V'(x_k)$ are negative along the entire trajectory;
- then make sure that the function $V(x_k)$ decreases at each step.
4. Results are presented as follows:
- phase trajectories are plotted;
- the total control effort is estimated;
- visualization of system behavior.

Distributed implementation

For a distributed implementation, each subsystem solves its own local optimization problem:

$$J_i(u_i) = \sum_{k=0}^{N-1} \left[V_i(x_i^k) + \rho \left\| x_i^k - \bar{x}_{N_i}^k \right\|^2 + \beta \left\| u_i^k \right\|^2 \right].$$

The coordination mechanism can be implemented in two ways:

1) the Jacobi method: at each iteration, the agent receives data from neighbors and recomputes its control;

2) ADMM: consensus variables and auxiliary Lagrange multipliers are introduced; simplification lies in the fact that the current implementation uses an approximate prototype without multipliers.

This approach provides:

- scalability;
- the ability to add or remove agents without changing the algorithm;
- stability even with incomplete convergence;
- feasibility of application in conditions of limited computing resources.

Empirical verification of the method

Verification objectives and structure

Empirical verification of the method is carried out to confirm its operability and stability in numerical modeling without linearization. The main tasks of verification are:

- confirmation of the correctness of the numerical solution;
- demonstration of a decreasing Lyapunov function along the trajectories;
- verification of consistency and stability in distributed configurations;
- visualization of system trajectories and qualitative interpretation of behavior.

The simulation is implemented using the Python programming language and the numpy, scipy and matplotlib modules. Verification includes two stages:

1) optimization of control on a single system;
2) distributed implementation with multiple agents.

Verification on a single system

System dynamics

The following nonlinear system is considered:

$$\begin{cases} \dot{x}_1 = -x_1 + x_2^2 + u \\ \dot{x}_2 = -x_2 + x_1 x_2 \end{cases}.$$

The Lyapunov function is chosen as:

$$V(x) = x_1^2 + x_2^2.$$

The initial state is set by the following parameters: $x_0 = [1{,}5; -1{,}0]$. Simulation time is $T = 10$, with a discretization of $N = 200$.

Without control

Without the control action $u(t) = 0$, the system exhibits unstable behavior. The Lyapunov function does not decrease and sometimes even increases. In the phase plane, the trajectory moves away from the origin.

The Python code to generate the plots (Figs. 1–3) is available for download and can be run locally[1].

Constant optimal control

The problem of finding $u = const$ that minimizes the objective functional for arbitrary control is solved:

$$J(u) = \alpha \int_0^T V\left(x(t)\right) dt + \beta \int_0^T u^2 dt.$$

Substituting the condition $u = const$ into this equation, we obtain:

$$J(u) = \alpha \int_0^T V\left(x(t;u)\right) dt + \beta \int_0^T u^2 dt = \alpha \int_0^T V\left(x(t;u)\right) dt + \beta T u^2,$$

where the trajectory $x(t; u)$ can be expressed by the system:

$$\begin{cases} \dot{x}_1 = -x_1 + x_2^2 + u \\ \dot{x}_2 = -x_2 + x_1 x_2 \\ x(0) + [1.5 - 1.0]^T \end{cases}.$$

At $\alpha = 1$, $\beta = 0{,}1$, the optimal value $u^* \approx -0{,}71$ ensures a guaranteed decrease in $V(x(t))$ and a bounded phase trajectory (the system decays without emissions).

[1] GitHub – fershtadt/lyapunov-control-visualization, Available: https://github.com/fershtadt/lyapunov-control-visualization (Accessed 13.11.2025).

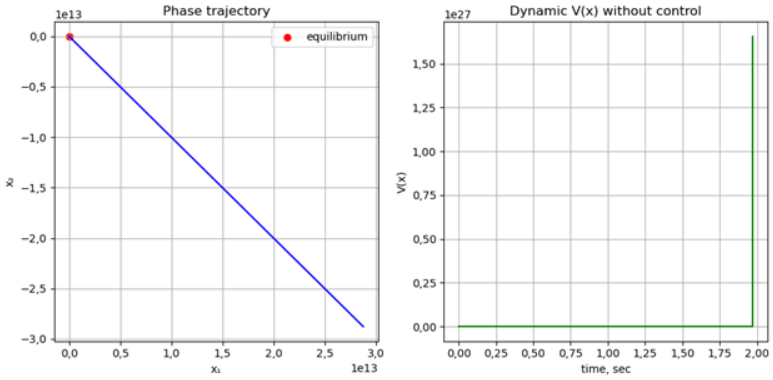


Fig. 1. Simulation of the system without control

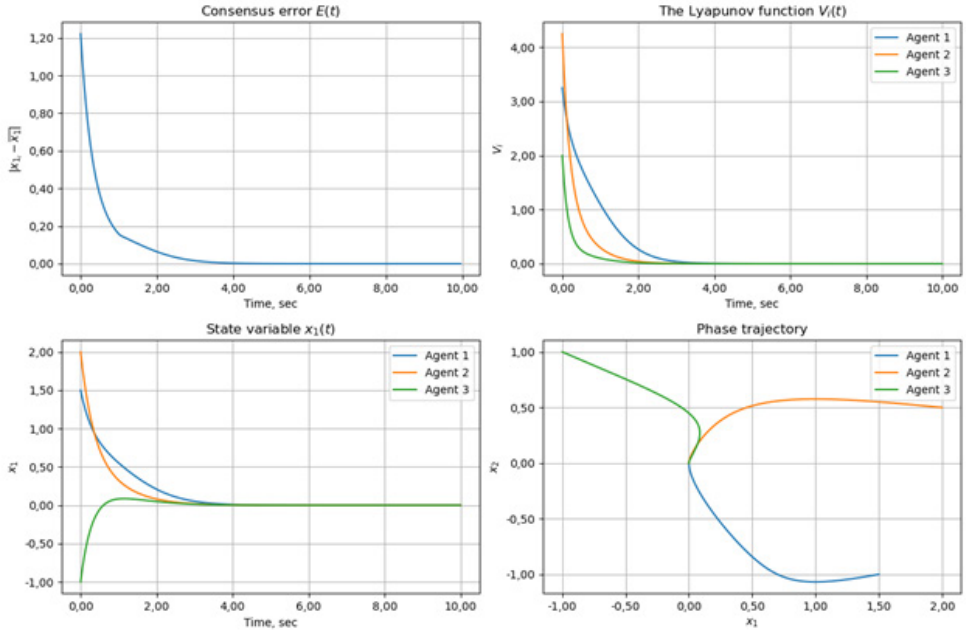


Fig. 2. Simulation with variable control

Variable control

For discrete control u_k ($k = 0, ..., N-1$), the following objective functional is minimized:

$$J(u) = \sum_{k=0}^{N-1} \left[\alpha V(x_k) + \beta u_k^2 \right] \Delta t.$$

Results:

- the Lyapunov function $V(x)$ decreases monotonically;
- trajectories $x_1(t)$, $x_2(t)$ tend to zero without oscillations;
- the derivative $V'(x) < 0$ along the entire trajectory;
- no overflow, oscillations, or anomalies are observed.
- the trajectories $V'(t)$ for each agent lie below zero, therefore, the condition $V' < 0$ along the trajectory is confirmed.

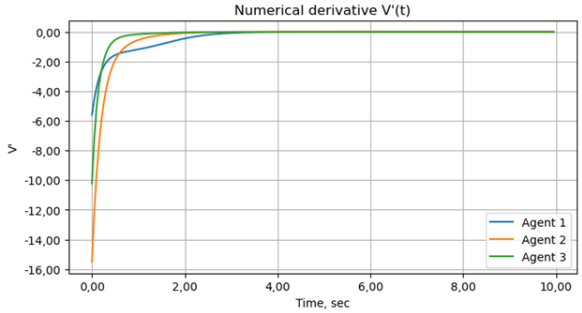


Fig. 3. Numerical derivative of $V(t)$

Conclusions

1. A constant optimal value of u^* stabilizes the system, but the attenuation of $V(t)$ is slower than with an optimized time-varying control strategy.

2. Optimizing the discrete control sequence u_k accelerates convergence and reduces control effort, as evidenced by the steeper decline of $V(t)$ and the negative derivative $V'(t)$ over the entire interval for each agent.

3. It has been empirically confirmed that including the Lyapunov function in the objective vector makes it possible to achieve stability without linearization, even with a simple Euler discretization and a small number of Jacobi coordination iterations.

Verification in a distributed system

System configuration

A system of three agents is being considered:

$$\dot{x}_i = f_i\left(x_i,\, u_i\right) + \sum_{j \in N_i} \gamma_{ij}\left(x_j - x_i\right),$$

where: N_i is the set of neighbors of agent i; $\gamma_{ij} = 0{,}5$ are consensus coefficients.

Each agent has the same local dynamics as the single-system and uses a local control of the form:

$$u_i^k = -Kx_{1,i}^k - \rho\left(x_{1,i}^k - \overline{x}_{1,N_i}^k\right).$$

Distributed optimization

A simulation of distributed coordination using the Jacobi method was implemented in Python (Figs. 2, 3). Each agent updates its trajectory according to its own equations, using the average states of its neighbors.

Jacobi iterations: 20.

Results:

- all agents demonstrate coherent behavior: deviations between them decrease;
- the Lyapunov function decreases for each agent;
- control inputs $u_i(t)$ remain within the acceptable range;
- the agents' phase trajectories converge to a common equilibrium state.

Conclusion

Even in the simplest implementation of distributed optimization (without Lagrange multipliers), the proposed approach provides:

- stability of the entire system;
- coordination between agents;
- numerical stability in nonlinear dynamics.

REFERENCES

1. **Shashikhin V.N., Budnik S.V., Golovina K.O.** Control of the spectrum of Lyapunov characteristic exponents in nonlinear large-scale systems. Computing, Telecommunication and Control, 2021, Vol. 14, No. 4, Pp. 37–51. DOI: 10.18721/JCSTCS.14404

2. **Shashikhin V.N., Savchuk S.A.** Complete synchronization of chaotic systems at different parameter values. Computing, Telecommunication and Control, 2022, Vol. 15, No. 2, Pp. 8–24. DOI: 10.18721/JCSTCS.15201

3. **Zhang C., Fu J.** Multi-objective dynamic optimization of path-constrained switched systems. Automatica, 2024, Vol. 159, Art. no. 111326. DOI: 10.1016/j.automatica.2023.111326

4. **Tao M., Li Q., Yu J.** Multi-objective dynamic path planning with multi-agent deep reinforcement learning. Journal of Marine Science and Engineering, 2024, Vol. 13, No. 1, Art. no 20. DOI: 10.3390/jmse13010020

5. **Lin J., He C., Tian Y., Pan L.** Variable reconstruction for evolutionary expensive large-scale multiobjective optimization and its application on aerodynamic design. IEEE/CAA Journal of Automatica Sinica, 2025, Vol. 12, No. 4, Pp. 719–733. DOI: 10.1109/JAS.2024.124947

6. **Zheng Y., Wang Y., Li S.** Adaptive control Lyapunov function based model predictive control for continuous nonlinear systems. International Journal of Robust and Nonlinear Control, 2022, Vol. 33, No. 2, Pp. 1254–1266. DOI: 10.1002/rnc.6409

7. **Zheng X.-Y., Yan H.-S.** Lyapunov-based stochastic model predictive control of stochastic nonlinear systems with input-delay using multi-dimensional Taylor network. Journal of the Franklin Institute, 2025, Vol. 362, No. 9, Art. no. 107694. DOI: 10.1016/j.jfranklin.2025.107694

8. **Baheri A.** Distributionally robust Lyapunov–Barrier Networks for safe and stable control under uncertainty. Results in Control and Optimization, 2025, Vol. 19, Art. no. 100556. DOI: 10.1016/j.rico.2025.100556

9. **Stiti C., Benrabah M., Aouaichia A., Oubelaid A., Bajaj M., Tuka M.B., Kara K.** Lyapunov-based neural network model predictive control using metaheuristic optimization approach. Scientific Reports, 2024, Vol. 14, Art. no. 18760. DOI: 10.1038/s41598-024-69365-9

10. **Nersesov S.G., Haddad W.M.** Control vector Lyapunov functions for large-scale impulsive dynamical systems. Nonlinear Analysis: Hybrid Systems, 2007, Vol. 1, No. 2, Pp. 223–243. DOI: 10.1016/j.nahs.2006.10.006

11. **Chen Y.-H.** Nonlinear adaptive optimal control design and implementation for trajectory tracking of four-wheeled mecanum mobile robots. Mathematics, 2024, Vol. 12, No. 24, Art. no. 4013. DOI: 10.3390/math12244013

12. **Yan Z., Zhang M., Zhou J., Yue L.** Distributed Lyapunov-based model predictive control for AUV formation systems with multiple constraints. Journal of Marine Science and Engineering, 2024, Vol. 12, No. 3, Art. no. 363. DOI: 10.3390/jmse12030363

13. **Peitz S., Dellnitz M.** A survey of recent trends in multiobjective optimal control – surrogate models, feedback control and objective reduction. Mathematical and Computational Applications, 2018, Vol. 23, No. 2, Art. no. 30. DOI: 10.3390/mca23020030

14. **Wen J., Li L., Wu Q., Li K., Lu J.** Multi-objective cooperative adaptive cruise control platooning of intelligent connected commercial vehicles in event-triggered conditions. Actuators, 2024, Vol. 13, No. 12, Art. no. 522. DOI: 10.3390/act13120522

15. **Ionescu C.M., Caruntu C.F., Cajo R., Ghita M., Crevecoeur G., Copot C.** Multi-objective predictive control optimization with varying term objectives: A wind farm case study. Processes, 2019, Vol. 7, No. 11, Art. no. 778. DOI: 10.3390/pr7110778

16. **El Sayed M.A., Farahat F.A., Elsisy M.A., Alsabaan M., Ibrahem M.I., Elwahsh H.** Two TOPSIS-based approaches for multi-choice rough bi-level multi-objective nonlinear programming problems. Mathematics, 2025, Vol. 13, No. 8, Art. no. 1242. DOI: 10.3390/math13081242

17. **Nelyubin A.P., Podinovski V.V.** Multicriteria problems with importance-ordered criteria groups. Automation and Remote Control, 2022, Vol. 83, No. 7, Pp. 1108–1122. DOI: 10.1134/S0005117922070074

18. **Podinovski V.V., Potapov M.A.** Analysis of the sensitivity of solutions of multi-criteria problems based on parametric partial preference relations. Automation and Remote Control, 2019, Vol. 80, No. 7, Pp. 1294–1303. DOI: 10.1134/S0005117919070075

19. **Balandin D.V., Kogan M.M.** Multicriteria robust generalized H_2 and γ_0 controllers with application to stabilization of a rotor in electromagnetic bearings. Automation and Remote Control, 2018, Vol. 79, No. 6, Pp. 996–1012. DOI: 10.1134/S0005117918060024

20. **Demidenko O.M., Borchik E.M., Yakimov A.I.** Multi-criterial optimization of resource distribution in the process of finished products. Problems of Physics, Mathematics and Technics, 2022, No. 3 (52), Pp. 90–96. DOI: 10.54341/20778708_2022_3_52_90

21. **Melnikov A.V., Karavaev A.A., Zhelezniakov A.O.** Analysis of the main approaches to modeling the management processes of organizational and technical systems for special purposes. Herald of Dagestan State Technical University. Technical Sciences, 2024, Vol. 51, No. 1, Pp. 153–165. DOI: 10.21822/2073-6185-2024-51-1-153-165

22. **Belykh M.A.** Formalization of a multi-criteria transport task with time constraints. Modeling, Optimization and Information Technology, 2024, Vol. 12, No. 2, Art. no. 27. DOI: 10.26102/2310-6018/2024.45.2.027

INFORMATION ABOUT AUTHORS / СВЕДЕНИЯ ОБ АВТОРАХ

Mikhail I. Fershtadt
Ферштадт Михаил Иосифович
E-mail: tral1930@mail.ru

Vladimir N. Shashikhin
Шашихин Владимир Николаевич
E-mail: shashihin@bk.ru

Submitted: 04.06.2025; Approved: 15.10.2025; Accepted: 16.11.2025.
Поступила: 04.06.2025; Одобрена: 15.10.2025; Принята: 16.11.2025.