



БЕЛОРУССКИЙ
ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ

ЖУРНАЛ
БЕЛОРУССКОГО ГОСУДАРСТВЕННОГО
УНИВЕРСИТЕТА

МАТЕМАТИКА ИНФОРМАТИКА

JOURNAL
OF THE BELARUSIAN STATE UNIVERSITY

MATHEMATICS and INFORMATICS

Издается с января 1969 г.
(до 2017 г. – под названием «Вестник БГУ.
Серия 1, Физика. Математика. Информатика»)

Выходит три раза в год

1

2026

МИНСК
БГУ

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

Главный редактор **ХАРИН Ю. С.** – доктор физико-математических наук, академик НАН Беларуси, профессор; директор Научно-исследовательского института прикладных проблем математики и информатики Белорусского государственного университета, Минск, Беларусь.
E-mail: kharin@bsu.by

**Заместители
главного редактора** **КРОТОВ В. Г.** – доктор физико-математических наук, профессор; профессор кафедры теории функций механико-математического факультета Белорусского государственного университета, Минск, Беларусь.
E-mail: krotov@bsu.by

ДУДИН А. Н. – доктор физико-математических наук, профессор; заведующий лабораторией прикладного вероятностного анализа факультета прикладной математики и информатики Белорусского государственного университета, Минск, Беларусь.
E-mail: dudin@bsu.by

**Ответственный
секретарь** **МАТЕЙКО О. М.** – кандидат физико-математических наук, доцент; доцент кафедры общей математики и информатики механико-математического факультета Белорусского государственного университета, Минск, Беларусь.
E-mail: matseika@bsu.by

- Абламейко С. В.* Белорусский государственный университет, Минск, Беларусь.
Альтенбах Х. Магдебургский университет им. Отто фон Герике, Магдебург, Германия.
Антоневич А. Б. Белорусский государственный университет, Минск, Беларусь.
Бауэр С. М. Санкт-Петербургский государственный университет, Санкт-Петербург, Россия.
Беняш-Кривец В. В. Белорусский государственный университет, Минск, Беларусь.
Босяков С. М. Белорусский государственный университет, Минск, Беларусь.
Бухштабер В. М. Математический институт им. В. А. Стеклова Российской академии наук, Московский государственный университет им. М. В. Ломоносова, Москва, Россия.
Вабищевич П. Н. Институт проблем безопасного развития атомной энергетики Российской академии наук, Москва, Россия.
Волков В. М. Белорусский государственный университет, Минск, Беларусь.
Гладков А. Л. Белорусский государственный университет, Минск, Беларусь.
Го В. Китайский университет науки и технологий, Хэфэй, провинция Аньхой, Китай.
Гогинава У. Тбилисский государственный университет им. Иванэ Джавахишвили, Тбилиси, Грузия.
Головко В. А. Брестский государственный технический университет, Брест, Беларусь.
Гороховик В. В. Институт математики Национальной академии наук Беларуси, Минск, Беларусь.
Громак В. И. Белорусский государственный университет, Минск, Беларусь.
Демид Г. Институт математики и информатики Вильнюсского университета, Вильнюс, Литва.
Егоров А. Д. Институт математики Национальной академии наук Беларуси, Минск, Беларусь.
Еремеев В. А. Гданьский политехнический университет, Гданьск, Польша.
Жоландек Х. Институт математики Варшавского университета, Варшава, Польша.
Журавков М. А. Белорусский государственный университет, Минск, Беларусь.
Залесский П. А. Бразильский университет, Бразилиа, Бразилия.
Каплунов Ю. Д. Университет Кииле, Кииле, Великобритания.
Кашин Б. С. Математический институт им. В. А. Стеклова Российской академии наук, Московский государственный университет им. М. В. Ломоносова, Москва, Россия.
Келлерер Х. Грацский университет им. Карла и Франца, Грац, Австрия.
Кожанов А. И. Институт математики им. С. Л. Соболева, Новосибирский государственный университет, Новосибирск, Россия.
Котов В. М. Белорусский государственный университет, Минск, Беларусь.
Краснопрошин В. В. Белорусский государственный университет, Минск, Беларусь.
Лауринчикас А. П. Вильнюсский университет, Вильнюс, Литва.

- Мадани К.** Университет Париж-Эст Марн-ла-Валле, Марн-ла-Валле, Франция.
- Макаров Е. К.** Институт математики Национальной академии наук Беларуси, Минск, Беларусь.
- Матус П. П.** Институт математики Национальной академии наук Беларуси, Минск, Беларусь.
- Медведев Д. Г.** Белорусский государственный университет, Минск, Беларусь.
- Михасев Г. И.** Харбинский политехнический университет, Харбин, Китай.
- Нестеренко Ю. В.** Московский государственный университет им. М. В. Ломоносова, Москва, Россия.
- Никоноров Ю. Г.** Южный математический институт Владикавказского научного центра Российской академии наук, Владикавказ, Россия.
- Освальд П.** Боннский университет, Бонн, Германия.
- Романовский В. Г.** Мариборский университет, Марибор, Словения.
- Сафонов В. Г.** Институт математики НАН Беларуси, Минск, Беларусь.
- Скиба А. Н.** Гомельский государственный университет им. Франциска Скорины, Гомель, Беларусь.
- Сотсков Ю. Н.** Объединенный институт проблем информатики Национальной академии наук Беларуси, Минск, Беларусь.
- Трофимов В. А.** Южно-Китайский университет технологий, Гуанчжоу, Китай.
- Тузинов А. В.** Объединенный институт проблем информатики Национальной академии наук Беларуси, Минск, Беларусь.
- Фильцмозер П.** Венский технический университет, Вена, Австрия.
- Черноусов В. И.** Альбертский университет, Эдмонтон, Канада.
- Чижик С. А.** Национальная академия наук Беларуси, Минск, Беларусь.
- Шешок Д.** Вильнюсский технический университет им. Гедиминаса, Вильнюс, Литва.
- Шубэ А. С.** Институт математики и информатики Академии наук Республики Молдова, Кишинев, Молдова.

EDITORIAL BOARD

- | | |
|------------------------------------|--|
| Editor-in-chief | KHARIN Y. S. , doctor of science (physics and mathematics), academician of the National Academy of Sciences of Belarus, full professor; director of the Research Institute for Applied Problems of Mathematics and Informatics, Belarusian State University, Minsk, Belarus.
E-mail: kharin@bsu.by |
| Deputy
editors-in-chief | KROTOV V. G. , doctor of science (physics and mathematics), full professor; professor at the department of function theory, faculty of mechanics and mathematics, Belarusian State University, Minsk, Belarus.
E-mail: krotov@bsu.by |
| | DUDIN A. N. , doctor of science (physics and mathematics), full professor; head of the laboratory of applied probabilistic analysis, faculty of applied mathematics and computer science, Belarusian State University, Minsk, Belarus.
E-mail: dudin@bsu.by |
| Executive
secretary | MATEIKO O. M. , PhD (physics and mathematics), docent; associate professor at the department of general mathematics and computer science, faculty of mechanics and mathematics, Belarusian State University, Minsk, Belarus.
E-mail: matseika@bsu.by |
-
- | | |
|------------------------------|---|
| Ablameyko S. V. | Belarusian State University, Minsk, Belarus. |
| Altenbach H. | Otto-von-Guericke University, Magdeburg, Germany. |
| Antonevich A. B. | Belarusian State University, Minsk, Belarus. |
| Bauer S. M. | Saint Petersburg State University, Saint Petersburg, Russia. |
| Beniash-Kryvets V. V. | Belarusian State University, Minsk, Belarus. |
| Bosyakov S. M. | Belarusian State University, Minsk, Belarus. |
| Buchstaber V. M. | Steklov Institute of Mathematics of Russian Academy of Sciences, Lomonosov Moscow State University, Moscow, Russia. |
| Chernousov V. I. | University of Alberta, Edmonton, Canada. |
| Chizhik S. A. | National Academy of Sciences of Belarus, Minsk, Belarus. |
| Dzemyda G. | Institute of Mathematics and Informatics of the Vilnius University, Vilnius, Lithuania. |
| Egorov A. D. | Institute of Mathematics of the National Academy of Sciences of Belarus, Minsk, Belarus. |
| Eremeyev V. A. | Gdansk University of Technology, Gdansk, Poland. |
| Filzmoser P. | Vienna University of Technology, Vienna, Austria. |
| Gladkov A. L. | Belarusian State University, Minsk, Belarus. |
| Goginava U. | Ivane Javakhishvili Tbilisi State University, Tbilisi, Georgia. |
| Golovko V. A. | Brest State Technical University, Brest, Belarus. |
| Gorokhovik V. V. | Institute of Mathematics of the National Academy of Sciences of Belarus, Minsk, Belarus. |
| Gromak V. I. | Belarusian State University, Minsk, Belarus. |
| Guo W. | University of Science and Technology of China, Hefei, Anhui, China. |
| Kaplunov J. D. | Keele University, Keele, United Kingdom. |
| Kashin B. S. | Steklov Institute of Mathematics of Russian Academy of Sciences, Lomonosov Moscow State University, Moscow, Russia. |
| Kellerer H. | University of Graz, Graz, Austria. |
| Kotov V. M. | Belarusian State University, Minsk, Belarus. |
| Kozhanov A. I. | Sobolev Institute of Mathematics, Novosibirsk State University, Novosibirsk, Russia. |
| Krasnoproshin V. V. | Belarusian State University, Minsk, Belarus. |
| Laurinchikas A. P. | Vilnius University, Vilnius, Lithuania. |
| Madani K. | Université Paris-Est, Marne-la-Vallée, France. |
| Makarov E. K. | Institute of Mathematics of the National Academy of Sciences of Belarus, Minsk, Belarus. |
| Matus P. P. | Institute of Mathematics of the National Academy of Sciences of Belarus, Minsk, Belarus. |
| Medvedev D. G. | Belarusian State University, Minsk, Belarus. |
| Mikhasev G. I. | Harbin Institute of Technology, Harbin, China. |
| Nesterenko Y. V. | Lomonosov Moscow State University, Moscow, Russia. |
| Nikonorov Y. G. | Southern Mathematical Institute of the Vladikavkaz Scientific Center of the Russian Academy of Sciences, Vladikavkaz, Russia. |

- Oswald P.** University of Bonn, Bonn, Germany.
- Romanovskij V. G.** University of Maribor, Maribor, Slovenia.
- Safonov V. G.** Institute of Mathematics of the National Academy of Sciences of Belarus, Minsk, Belarus.
- Šešok D.** Vilnius Gediminas Technical University, Vilnius, Lithuania.
- Skiba A. N.** Francisk Skorina Gomel State University, Gomel, Belarus.
- Sotskov Y. N.** United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus.
- Suba A. S.** Institute of Mathematics and Computer Science of the Academy of Sciences of Moldova, Kishinev, Moldova.
- Trofimov V. A.** South China University of Technology, Guangzhou, China.
- Tuzikov A. V.** Research Institute for Applied Problems of Mathematics and Informatics of the National Academy of Sciences of Belarus, Minsk, Belarus.
- Vabishchevich P. N.** Institute for the Safe Development of Atomic Energy of the Russian Academy of Sciences, Moscow, Russia.
- Volkov V. M.** Belarusian State University, Minsk, Belarus.
- Zalesskii P. A.** University of Brazilia, Brazilia, Brazil.
- Zhuravkov M. A.** Belarusian State University, Minsk, Belarus.
- Zoladek H.** Mathematics Institute of the University of Warsaw, Warsaw, Poland.

КОМПЬЮТЕРНЫЙ АНАЛИЗ ДАННЫХ

COMPUTER DATA ANALYSIS

УДК 519.2:004.7

СОВРЕМЕННЫЕ ТЕНДЕНЦИИ В КОМПЬЮТЕРНОМ АНАЛИЗЕ ДАННЫХ

Ю. С. ХАРИН¹⁾, С. В. АБЛАМЕЙКО²⁾

¹⁾Научно-исследовательский институт прикладных проблем математики и информатики БГУ,
пр. Независимости, 4, 220030, г. Минск, Беларусь

²⁾Белорусский государственный университет, пр. Независимости, 4, 220030, г. Минск, Беларусь

Аннотация. Описываются современные тенденции в компьютерном анализе данных, представленные на конференции «Компьютерный анализ данных и моделирование» и конгрессе «Информационные системы и технологии».

Ключевые слова: статистический анализ данных; робастность; стохастика; дискретные данные; нейронные сети; компьютерное зрение.

Образец цитирования:

Харин ЮС, Абламейко СВ. Современные тенденции в компьютерном анализе данных. *Журнал Белорусского государственного университета. Математика. Информатика*. 2026; 1:6–15 (на англ.).
EDN: EHSTDO

For citation:

Kharin YuS, Ablameyko SV. Current trends in computer data analysis. *Journal of the Belarusian State University. Mathematics and Informatics*. 2026;1:6–15.
EDN: EHSTDO

Авторы:

Юрий Семенович Харин – доктор физико-математических наук, академик НАН Беларуси, профессор; директор.
Сергей Владимирович Абламейко – доктор технических наук, академик НАН Беларуси, профессор; профессор кафедры веб-технологий и компьютерного моделирования механико-математического факультета.

Authors:

Yuriy S. Kharin, doctor of science (physics and mathematics), academician of the National Academy of Sciences of Belarus, full professor; director.
kharin@bsu.by
<https://orcid.org/0000-0003-4226-2546>
Sergey V. Ablameyko, doctor of science (engineering), academician of the National Academy of Sciences of Belarus, full professor; professor at the department of web technologies and computer simulation, faculty of mechanics and mathematics.
ablameyko@bsu.by
<https://orcid.org/0000-0001-9404-1206>

CURRENT TRENDS IN COMPUTER DATA ANALYSIS

Yu. S. KHARIN^a, S. V. ABLAMEYKO^b

^aResearch Institute for Applied Problems of Mathematics and Informatics, Belarusian State University,
4 Niezaliezhnasci Avenue, Minsk 220030, Belarus

^bBelarusian State University, 4 Niezaliezhnasci Avenue, Minsk 220030, Belarus

Corresponding author: Yu. S. Kharin (kharin@bsu.by)

Abstract. The current trends in computer data analysis presented at the conference «Computer data analysis and modelling: stochastics and data science» and the congress «Information systems and technologies» are described.

Keywords: statistical data analysis; robustness; stochastics; count data; neural networks; computer vision.

Introduction

The special issue «Stochastics and data science» of the publication «Journal of Belarusian State University. Mathematics and Informatics» is based on extended papers presented at the 14th International scientific conference «Computer data analysis and modelling: stochastics and data science» and the 11th International congress on computer science «Information systems and technologies», which were held in 2025 in Minsk. All articles underwent the double-blind peer review process standard for the journal.

The structure of this paper is as follows: sections «International scientific conference “Computer data analysis and modelling”» and «International congress on computer science “Information systems and technologies”» are devoted to short description of the indicated conference and congress; sections «Current trends in statistical data analysis» and «Current trends in data science» are presented our version of current trends in computer data analysis and in data science, respectively; section «Short description of the contributed papers» presents short descriptions of the contributed papers.

International scientific conference «Computer data analysis and modelling»

International scientific conference «Computer data analysis and modelling» (CDAM) has been held regularly (every two or three years) since 1988. This conference was held in September 1988, December 1990, December 1992, September 1995, June 1998, September 2001, September 2004, September 2007, September 2010, September 2013, September 2016, September 2019, September 2022 and September 2025 in Minsk.

The CDM conference is currently the largest scientific forum on statistical data analysis in the post-Soviet area attracting significant interest both in Belarus and the CIS countries, as well as in countries further abroad. Nowadays the conference is included into the list of conferences of the International Institute of Mathematical Statistics. The goal of the CDM conference is to organise a broad scientific forum for researchers from various countries to discuss and to exchange the latest scientific results in the field of data analysis theory, statistical modelling, statistical software, as well as their applications in scientific research, economics, information security, medicine and in other areas.

This conference is organised by the Belarusian State University (BSU), the Research Institute for Applied Problems of Mathematics and Informatics of BSU, Lomonosov Moscow State University and the Vienna University of Technology. Traditionally, the conferences are held with the financial support of the Ministry of Education of the Republic of Belarus, the Belarusian Republican Foundation for Fundamental Research and the Scientific and Technological Association «Infopark». Holding such a forum at BSU is not random. BSU is the country's leading scientific centre in a set of priority research areas related to the development and use of new information technologies. This includes computer data analysis and modelling of complex stochastic systems in various fields: scientific research, digital economy, manufacturing, information security, medicine, banking, business and other. Also, the faculty of applied mathematics and computer science of BSU trains bachelor's degree specialists to work in these fields under the following programmes: applied mathematics, computer science, applied informatics and computer security. Additionally, the master's programme in computer data analysis is offered.

The previous conference in 2022 was attended by over 100 scientists and young researchers from Austria, Belarus, Germany, Italy, Kazakhstan, Russia, the USA, Uzbekistan, the Czech Republic and Luxembourg. The programme and organising committee members were leading experts in the field of robust and computer methods of data analysis and stochastic systems modelling from seven countries.

The theme of the CDAM conference that held in 2025 is «Stochastics and data analysis» (fig. 1). It combines fundamental science, namely mathematical statistics as the theoretical foundation of data science, with relevant applied research. This theme encompasses mathematical models, methods, algorithms and software for data analysis and stochastic modelling. The goal is to obtain optimal statistical inferences (decisions, estimates, forecasts), as well as to develop computer models of complex stochastic systems and processes. This area is classified among the priority research directions approved by the Council of Ministers of the Republic of Belarus.



Fig. 1. Participants of the CDAM conference that held in 2025

The conference significance is also underscored by the participation of young scientists, master's degree students, postgraduate students and undergraduate students. As it is well known, the majority of winners of various-level school, university and international olympiads in mathematics and computer science successfully study at the faculty of applied mathematics and computer science of BSU. Starting from their 1st or 2nd year, they are engaged in research within student scientific clubs, student research laboratories and the Research Institute for Applied Problems of Mathematics and Informatics of BSU.

Examples of important applied achievements by Belarusian scientists include:

- development and testing of computer systems for information security in the banking sector of the Republic of Belarus;
- creation of econometric models and leading economic indicators for analysing and forecasting the dynamics of key macroeconomic and monetary indicators, as well as algorithms and software for the statistical analysis of the economic creditworthiness and the stability of the banking system at macro and micro levels;
- development of statistical methods and instrumental tools for solving applied problems in medicine for the N. N. Aleksandrov National Cancer Centre of Belarus, the Scientific and Practical Centre «Cardiology» and the department of medical radiology of the Ministry of Health of the Republic of Belarus.

International congress on computer science «Information systems and technologies»

Since 2002, BSU organised and began hosting the International scientific conference «Information systems and technologies». This conference was held every two years (2002, 2004, 2006, 2008, 2009, 2011, 2013, 2016, 2020 and 2022). It has become the International congress on computer science «Information systems and technologies» (CSIST) since 2011. The ideologist and main organiser of the conference throughout its existence was the dean of the faculty of applied mathematics and informatics of BSU P. A. Mandrik, and the faculty of applied mathematics and informatics of BSU served as its venue. CSIST is a large-scale and significant forum in the information technologies field. It aims to exchange experiences among the global scientific community, accumulate ideas and new solutions in information security and develop effective responses to modern challenges.

Let us note that CSIST that held in 2025 took place in 29–31 October 2025 in Minsk (fig. 2). The forum was organised by BSU and the United Institute of Informatics Problems of the National Academy of Sciences of Belarus. Over 200 scientists from Azerbaijan, Belarus, Kazakhstan, China, Portugal, Russia, Serbia, Turkey and Switzerland participated in the congress. It was opened on 30 October at BSU with a plenary session, featuring the most significant presentations by leading scientists. In particular, experts discussed computer vision (CV) in multimodal artificial intelligence (AI) systems, the use of deep learning to classify sleep stages and the development of an intelligent system for analysing patient medical data. The researchers also presented data from an analysis of human behaviour, when interacting with virtual environments, discussed topics such

as teacher training in the context of digital transformation, challenges of higher education during the information technology revolution and more.



Fig. 2. Opening ceremony of the CSIST that held in 2025

A total of 174 papers were submitted for the CSIST, and 124 articles were accepted for publication. The congress proceedings were published in two parts and comprise nearly 1 thsd pages. The number of paper authors was 237, including 35 doctors of science and 78 PhD.

Current trends in statistical data analysis

Robustness in statistical data analysis. The final result of statistical data analysis is a statistical inference (estimate, forecast, decision, etc.) based on some observed (registered) statistical data X . Respectively, statistical data analysis consists of different branches: statistical estimation, statistical forecasting (prediction), statistical classification (pattern recognition), etc. To be more precise let us describe the current trend «robustness» in connection with forecasting of future states in time series.

The mathematical substance of the statistical forecasting problem is quite simple – to estimate the future value $x_{T+\tau} \in \mathbb{R}^d$ of the d -variate time series in $\tau \in \mathbb{N}$ steps ahead by successive observations $X = (x_1, x_2, \dots, x_T) \in \mathbb{R}^{Td}$. In practice it was detected by many researchers that the optimal forecasting algorithms on the real statistical data have the risk values that are much more than the expected theoretical values. In lecture at the International congress of mathematicians P. Huber explained the reason of this strange situation as follows: «Statistical inferences (including statistical forecasts) depend only in part upon the observations. An equally important base is formed by prior assumptions about the underlying situation». The system of prior assumptions is called the hypothetical model of the data M_0 . In applied problems the hypothetical model assumptions M_0 are often distorted, and this fact leads to the instability of the optimal forecasting statistics that are optimal only under M_0 .

The probability model of the observed time series under distortions is determined by the family of probability distributions $\{P_T^\varepsilon(X) : X \in \mathbb{R}^{Td} : T \in \mathbb{N}, \varepsilon \in [0, \varepsilon_+]\}$, where ε is the distortion level; $\varepsilon_+ \geq 0$ is its maximal admissible value. If $\varepsilon_+ = 0$, then the distortions are absent, and we have the hypothetical model M_0 . A classification scheme for typical distortions of hypothetical models is given in work [1].

Let $\hat{x}_{T+\tau} = f(X) : \mathbb{R}^{Td} \rightarrow \mathbb{R}^d$ be any forecasting statistic. Its performance is evaluated by the mean square risk of forecasting

$$r_\varepsilon = r_\varepsilon(f) = \mathbb{E}_\varepsilon \left\{ \left\| \hat{x}_{T+\tau} - x_{T+\tau} \right\|^2 \right\} \geq 0, \quad \varepsilon \in [0, \varepsilon_+],$$

where $\mathbb{E}_\varepsilon \{\cdot\}$ is the expectation symbol with reference to the probability distribution $P_T^\varepsilon(\cdot)$. If the distortions are absent ($\varepsilon = 0$) and the hypothetical model M_0 is valid, this functional is called the hypothetical risk $r_0 = r_0(f)$. The guaranteed (upper) risk is

$$r_+ = r_+(f) = \sup_{0 \leq \varepsilon \leq \varepsilon_+} r_\varepsilon(f),$$

where the supremum is taken on all admissible distortions of the hypothetical model M_0 . Let $\hat{x}_{T+\tau}^0 = f^0(X)$ be the optimal forecasting statistic under the known hypothetical model M_0 ($\varepsilon = 0$) that gives the minimal value to the hypothetical risk

$$r_0 = r_0(f^0) = \inf_{f(\cdot)} r_0(f).$$

The risk instability coefficient k is the relative increment of the guaranteed risk with reference to the minimal hypothetical risk:

$$k = k(f) = \frac{r_+(f) - r_0}{r_0 \geq 0}.$$

For any $\delta > 0$ define a characteristic of robustness that is called the δ -admissible distortion level and is quite useful in applications:

$$\varepsilon^* = \varepsilon^*(\delta) = \sup \{ \varepsilon \in [0, \varepsilon_+] : k(f) \leq \delta \}.$$

It indicates the maximal level of distortions for which the relative increment of the risk is yet not greater than the fixed value $\delta \cdot 100\%$.

The smaller the value of k and the greater the value of ε^* , the more robust the forecasting statistic. The minimax robust forecasting statistic $\hat{x}_{T+\tau}^* = f^*(X)$ minimises the risk instability coefficient:

$$k(f^*) = \inf_{f(\cdot)} k(f).$$

Further expansion of robustness paradigm in statistical forecasting is possible in two main directions [1–9]: 1) using new distortion types for the underlying hypothetical model of observed data; 2) using new more complex hypothetical models.

Considering the first direction, let us indicate some distortion types that are topical for practice:

- presence of incomplete data including, namely missing values (when a portion of d_t , $1 \leq d_t \leq d$, components of x_t at the time point t is not observed), censoring (when at the time point t not exact value x_t but only an event $\{x_t \in A_t\}$, $A_t \subset \mathbb{R}^d$, concerning x_t is observed); e. g. $A_t = \{x_{t,1} > c\}$, where $c \in \mathbb{R}^1$ is some known threshold;
- simultaneous (joint) influence of several distortion types on the underlying hypothetical model, e. g. simultaneous influence of outliers and jointly missing values are present.

Concerning the second direction in the development of robust forecasting, let us mention some complex (non-conventional) hypothetical probabilistic models actively used in data analysis:

- multivariate discrete-valued time series $x_t = (x_{t,i}) \in V^d$, where V is some discrete set of cardinality $2 \leq |V| \leq +\infty$ (this model is extremely helpful in many applied areas, e. g. in genetics ($d = 1$, $|V| = 4$), information protection ($d = 64$, $|V| = 2$));
- spatial or spatio-temporal data (random fields) $x_t \in V^d$, where $t = (t_j) \in \mathbb{R}^N$ is an N -dimensional parameter indicating spatial or jointly spatial and temporal coordinates of an observation x_t , e. g. images, meteorological maps, dynamical maps of incidence rates.

Statistical analysis of discrete-valued time series. In any digitised society a multitude of discrete-valued data is generated. If these discrete-valued data are considered dynamically (as dependent on discrete time $t \in \mathbb{Z}$), we obtain a discrete-valued time series $x_t \in A$, where A is a discrete set. In particular, if $A = \{0, \dots, N-1\}$, $N = |A|$, $2 \leq N \leq +\infty$, we obtain a count time series.

The theory of statistical analysis of time series is deeply developed only for the so-called continuous time series, where A is a subset of $\mathbb{R}^d$ with non-zero Lebesgue measure. For the discrete case with discrete-valued observations the theory is only in the beginning stage. The main theoretical problem here is modelling of large depth stochastic dependence in discrete time series $\{x_t\}$.

A universal model for describing high-order stochastic dependencies is the homogeneous Markov chain (MC(s)) of sufficiently large order $s \in \mathbb{N}$ proposed by J. Doob and E. Dynkin, defined on some probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and specified by a generating equation for the conditional probability distribution

$$\mathbf{P}\{x_t = j_t | \mathcal{F}_{t-1}\} = \mathbf{P}\{x_t = j_t | X_{t-s}^{t-1} = J_{t-s}^{t-1}\} = p_{J_{t-s}^{t-1}, j_t}, \quad t \in \mathbb{Z},$$

where $\mathcal{F}_{t-1} = \sigma(\{x_\tau : \tau \leq t-1\})$ is the σ -algebra; $X_{t-s}^{t-1} = (x_{t-s}, \dots, x_{t-1})' \in A^s$; $J_{t-s}^{t-1} = (j_{t-s}, \dots, j_{t-1})' \in A^s$.

This model is defined by a matrix parameter – $(s+1)$ -dimensional transition probability matrix $P = p_{J_1^{s+1}, J_1^{s+1}} \in A^{s+1}$. The number of independent parameters for MC(s) ($D_{MC(s)} = N^s(N-1) = \mathcal{O}(N^{s+1})$) increases exponentially with the order s . To overcome the «curse of dimensionality» we propose to use parsimonious models (PMC(s)) for which the matrix P has the following parametric form:

$$p_{J_1^{s+1}} =: p_\alpha(J_1^{s+1}), \quad \alpha = (\alpha_1, \dots, \alpha_m)' \in \mathbb{R}^m,$$

where α is a low-dimensional ($m = D_{\text{PMC}(s)} \ll D_{\text{MC}(s)}$) parameter.

We propose four main approaches for constructing high-order parsimonious Markov chains: 1) the compression of the set of possible values for the elements of matrix P ; 2) the use of parametric families of discrete probability distributions; 3) the use of artificial neural networks (ANN) to approximate the dependence on past history; 4) an approach based on sufficient statistics and information geometry. Based on these approaches, we have managed to construct the following new parsimonious probabilistic models for discrete time series [10–12]:

- an s -order Markov chain with r partial connections $\text{MC}(s, r)$;
- a Markov chain of conditional order $\text{MCCO}(s, L)$;
- a binary conditionally non-linear autoregressive model $\text{BCNAR}(s)$;
- a binomial conditionally non-linear autoregressive model $\text{BiCNAR}(s)$;
- a semi-binomial conditionally non-linear autoregressive model $\text{SBiCNAR}(s)$;
- a Poisson conditionally non-linear autoregressive model $\text{PCNAR}(s)$;
- a neural network based model;
- a family of parsimonious models based on exponential families of probability distributions and sufficient statistics.

We have developed the frequencies based estimation method for statistical estimation of parameters for the constructed parsimonious models based on conditional multidimensional frequencies. This method allows for the construction of consistent statistical estimators and predictive statistics in an explicit form.

The directions in future development of statistical data analysis for discrete valued time series include:

- development of models, methods and software for multivariate discrete valued time series;
- formation of models, methods and software for multivariate discrete valued random fields;
- creation of mathematical theory and software for high-order Markov chains with time-varying transition matrices;
- development of stochastic volatility models.

Current trends in joint application of statistical data analysis methods and ANN. A necessary factor for the digitalisation of the economy is the processing of massive arrays of heterogeneous information with exponentially growing volumes. Currently, the universal approach considered as a «panacea» for big data is the approach based on the use of machine learning and ANN. This approach involves training an ANN (structure is selected empirically) on a training sample of a sufficiently large size. However, it is important to note three drawbacks of this empirical approach: 1) the lack of theoretical accuracy estimates for statistical inferences; 2) the absence of theoretical estimates for the required size of the training sample; 3) the lack of opportunity for the successful solution of problems in the statistical analysis of complex-structured data.

In our view, to overcome these shortcomings it is necessary to account for the fact that when constructing a solution d (an estimate, forecast or inference) from the available data X , the available prior information I must also be used: $d = f(X, I)$. The following situations are possible in this regard. If the probabilistic model, generating the data, is known, it is advisable to use classical methods from statistical decision theory. If the model is specified up to a parameter $\theta \in \Theta$, i. e. $I = M_\theta$, then methods of parametric statistics are applied. If the model uncertainty is even higher and it is defined by a functional set of probability distributions, methods of semi-parametric or non-parametric statistics can be employed. If prior information I is absent, one must be forced to use the ANN. It is worth note that even «minor» prior information can be useful in selecting the architecture and hyperparameters of the ANN. For example, to solve the problem in detection of statistical periodicity in a time series it is recommended to apply convolutional neural networks (CNN) with window sizes determined by the possible values of the period.

Current trends in data science

There are two main types of data depending on the way of their organisation: 1) structured data; 2) unstructured data. Structured data is information organised in a predefined, fixed format, typically in tables with rows and columns, making it easily searchable, processable and analysable by machines and humans. Unstructured data is information that lacks a predefined model or organisation, existing in native formats like text documents, images, audio, video and social media, making it hard for traditional systems to process but rich with valuable context, requiring AI and machine learning for analysis.

We will consider the second type of data. Unstructured data science involves extracting valuable insights from data that lacks a predefined format, like text, images, audio and video, using AI or machine learning techniques (natural language processing, computer vision) to find patterns, understand sentiment and fuel advanced analytics, making sense of the vast majority of enterprise data that does not fit traditional databases. Let us analyse the trends in unstructured data science.

Spiking neural networks as a third generation of neural networks. CNN currently dominate in image processing architectures. While lightweight CNN demonstrate moderate performance, they exhibit limited capacity to capture complex patterns and contextual dependencies essential for advanced tasks, thereby restricting their efficacy in practical CV scenarios. Conversely, while transformers effectively model complex relationships, their substantial computational overhead impedes compliance with computer constraints.

To overcome these limitations spiking neural networks (SNN) inspired by biological neuronal mechanisms represent a promising approach for image processing. SNN leverage event-driven dynamics and spatiotemporal encoding capabilities, combining biological plausibility with computational efficiency advantages suited to ultra-low-power environments. Although CNN have advanced image processing, SNN demonstrate superior potential in power-constrained scenarios. As a new generation of neural networks, SNN demonstrate the advantages of ultra-low power consumption and efficient computing through event-driven sparse calculations and biologically plausible spike transmission mechanisms, making them particularly suitable for image processing scenarios.

Brain-like computing can achieve architectural transitions through neuromorphic hardware. Brain-like intelligent computing has the computing advantages of low power consumption and high energy efficiency by imitating the computing architecture and methods of the human brain, providing a potential path to achieve the above technological breakthroughs. Compared with the second-generation ANN, the third-generation brain-like SNN has great potential advantages in high energy efficiency, high precision and high interpretability of remote sensing image intelligent processing due to its high bionics. Its structure is shown in fig. 3.

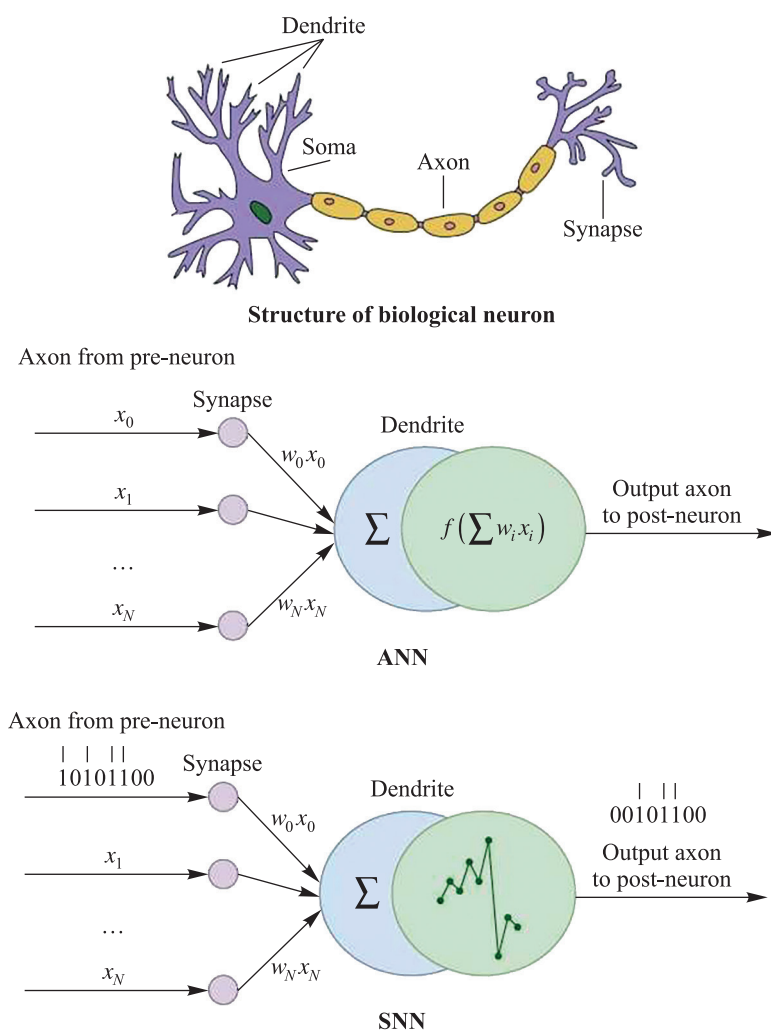


Fig. 3. ANN and SNN structure diagram

In 2019 SNN Spiking-YOLO for object detection tasks has been proposed [13]. Based on the YOLO network framework, it is improved by pulsing, which can extract multi-scale features. At the same time targeted structural changes are made to be applicable to input and output of different numbers of channels or dimensions. In 2024 X. Luo and his colleagues proposed the SpikeYOLO architecture, which combines the macrodesign of YOLOv8 and the microdesign of Meta-SpikeFormer [14]. It significantly narrows the performance gap between

SNN and ANN. Under similar parameters the performance of SpikeYOLO and YOLOv5 is comparable, and the energy efficiency is improved by 3.3 times.

SpikeYOLO is a breakthrough fusion model of SNN and a single-stage object detection framework. Its core goal is to introduce the event-driven mechanism of biological neurons into object recognition tasks. This architecture inherits the macroscopic detection paradigm of the YOLO series (such as feature pyramid and grid prediction mechanism) but replaces the convolution operation of the traditional ANN with a spiking neuron computing unit to achieve event-driven sparse computing. SpikeYOLO combines the macrodesign of YOLO with the microdesign of SNN modules.

The future trend of neural network development will be extending SNN for solving concrete CV tasks. For example, for object detection in remote sensing images.

Multi-modal data analysis. Multi-modal models refer to machine learning models capable of processing and integrating information from multiple modalities or types of data. These modalities can include text, images, audio, video and other forms of sensory input. Unlike traditional large language models that are typically designed to handle a single type of data, multi-modal models combine and analyse different forms of data inputs to achieve more comprehensive understanding and generate more robust outputs. As an example, a multi-modal model can receive a photo of a landscape as an input and generate a written summary of that place's characteristics. Also, it could receive a written summary of a landscape and generate an image based on that description. This ability to work across multiple modalities gives these models powerful capabilities.

In November 2022 OpenAI launched ChatGPT, which quickly put generative AI on the map. ChatGPT was unimodal AI designed to receive text inputs and generate text outputs by using natural language processing. GPT-4o introduced multi-modal capabilities to ChatGPT.

Key components of large multi-modal models are:

- data integration (LLMs use sophisticated algorithms to combine data from multiple sources, ensuring that the information from each modality is accurately represented and integrated);
- feature extraction (the model extracts relevant features from each type of input; for example, it might identify objects and their relationships in an image while understanding the context and meaning of accompanying text);
- joint representation (by creating a joint representation of the multi-modal data, the model can make inferences and generate outputs that consider all available information);
- cross-modal attention (techniques like cross-modal attention help the model to focus on relevant parts of the data from different modalities, improving its ability to generate coherent and contextually appropriate responses).

Currently, transformer-based multi-modal image recognition methods have shown remarkable performance in cross-modal information fusion, particularly by leveraging self-attention mechanisms to enhance interactions across modalities. Traditional multi-modal image recognition methods can be broadly categorised into three types: 1) early fusion; 2) late fusion; 3) hybrid fusion. Early fusion methods concatenate feature vectors from different modalities before classification [15]. While this approach improves the models accuracy to some extent, it ignores inherent differences in sampling rates and spatiotemporal resolutions across modalities, leading to temporal misalignment and feature heterogeneity issues [16]. Late fusion methods train separate unimodal classifiers and aggregate predictions via voting or stacking. Although this approach captures modality-specific features well, it fails to effectively capture the complementary relationships and contextual interactions between modalities, thus missing cross-modal synergistic discriminative information [17]. Hybrid fusion attempts to balance both strategies through multi-stage aggregation but still relies on static weights, unable to adapt to dynamic modality reliability, resulting in sharp performance drops in noisy environments or occluded scenarios [18].

Despite the significant progress in multi-modal image recognition, three key challenges persist: 1) reducing information loss while handling dynamic modality reliability; 2) preserving fine-grained emotional cues through multi-scale representations; 3) preventing representation drift during cross-modal semantic alignment. Traditional multi-modal image recognition methods often rely on cascaded CNN or recurrent neural networks to extract features from facial expressions, vocal cues and other modalities. However, these methods exhibit notable limitations in addressing dynamic modality reliability, multi-scale representation learning and cross-modal semantic alignment.

Multi-modal image recognition has gained significant attention due to advances in deep learning and availability of large multi-modal datasets. Despite these advancements, designing effective strategies for fusing and processing data from diverse modalities remains challenging.

Multi-modal learning is another new frontier. Traditional CV models rely on labeled data but labeling is costly and time-consuming. The move toward self-supervised learning allows models to learn from unlabeled data, using context to infer meaning. In future self-supervised learning will dominate, enabling faster model training, reduced dependency on human annotation and better generalisation across visual tasks.

CV and AI. AI is a multi-modal system that has the ability to process and understand inputs from different modes or modalities. While not dismissing the importance of advancements in language models and

acknowledging the significant progress in natural language processing, as exemplified by models like GPT, it is crucial to highlight the importance of CV for AI in understanding the physical world.

Computer perception is the ability of a machine to interpret and understand visual data from the world. CV in artificial general intelligence involves the development of algorithms and models that can process and interpret visual data, such as images and videos, to extract meaningful information. These algorithms can be trained to recognise objects, people and scenes, and to understand their spatial relationships and context. With this ability AI systems can analyse and interpret visual information to make informed decisions and take appropriate actions.

Also, CV is a crucial component in AI development. Computer perception is essential for an artificial general intelligence system to interact with the environment, understand the context of its surroundings and make intelligent decisions based on that understanding.

Despite the significant progress made in computer vision, developing an AI system with human-like perception abilities remains a challenge. One of the main challenges is the ability to integrate different sensory inputs seamlessly. For example, integrating visual and auditory information to understand a scene requires complex algorithms and machine learning models that can handle the high-dimensional data from various sensory modalities.

Deep learning has been instrumental in advancing the field of perception in AI development. Deep learning models, such as CNN and recurrent neural networks, have been used to extract features from sensory inputs and learn representations of the environment. For example, CNN have been used for object recognition tasks, while recurrent neural networks have been used for speech recognition tasks.

Computer perception plays a critical role in various AI applications, such as robotics, autonomous vehicles and healthcare. For example, an AI system with perception abilities can enable robots to navigate and interact with the environment autonomously, without human intervention. In healthcare an AI system with perception abilities can enable the early detection of diseases and assist doctors in making accurate diagnoses.

However, there are also challenges associated with perception in intelligent machines. One of the main challenges is that machines perceive the world differently than humans do. For example, machines may not be able to perceive emotions in the same way that humans do. This can make it difficult for machines to interact with humans in a natural way. Additionally, machines may not be able to perceive certain types of information, such as abstract concepts or social cues.

To overcome these challenges researchers are developing new approaches to perception in intelligent machines. One approach is to use machine learning to train machines to recognise and interpret different types of input. For example, machine learning algorithms can be used to teach machines to recognise different objects in an image. Another approach is to use multiple sensors to provide machines with a more complete view of their environment. For example, self-driving cars use a combination of cameras, lidar and radar to perceive their environment.

The future of perception in AI presents both opportunities and challenges. While machines with advanced perception capabilities will be able to perform tasks that were previously impossible, there are also challenges associated with building machines that perceive the world in the same way that humans do. However, researchers are making progress in the developing new approaches to perception in machines, and it is likely that we will continue to see advances in this field in the coming years.

Short description of the contributed papers

The introductory paper of Yu. S. Kharin and S. V. Ablameyko presents current trends in statistical data analysis and information technologies. In the article of M. P. Savelov asymptotic non-negative correlatedness of some statistics used in testing the quality of random number generators is analysed. The work of A. A. Serov is devoted to reliability of two-level testing approach of the NIST SP800-22 test suite. An integer random walk $\{S_i, i \geq 0\}$ with zero drift and finite variance σ^2 is considered, and the limit theorem on convergence to the local time of a Brownian excursion is presented in the publication of V. I. Afanasyev. The article of S. E. Vorobeychikov and Yu. B. Burkatovskaya is devoted to the problem of parameter estimation of the second order Markov modulated Poisson process and a new two-staged algorithm development. The work of V. A. Voloshko considers application of the local information geometry for statistical analysis of high order binary Markov chains. In the paper of A. Yu. Kharin a problem of sequential statistical hypotheses testing for discrete random data with block structure is considered. Y. L. Orlovich and M. A. Shutro present results on a topical problem on recognition of special classes of graphs. In the article of Y. R. Adamovskiy and R. P. Bohush smoke and camera tampering detection system based on single-board computer for real-time fire monitoring is presented. The work of K. V. Andrenko, A. A. Kroshchanka, V. A. Golovko and Yu. B. Krapivin is devoted to managerial decision analysis using small language models. Real-time gait recognition system based on lightweight dual-modal pose estimation is presented in the paper of Din Aodi, Ye Shiping, A. M. Nedzved and V. S. Anosov. The paper of V. V. Krasnoproshein

and A. A. Starovoitov presents a multi-model approach with vector search in operational management of critical IT service. Detection of image objects for autonomous localisation of unmanned aerial vehicles and map formation of the surrounding space is considered by V. Yu. Tsviatkou, V. V. Rabtsevich, Ma Jun and M. I. Zorko.

References

1. Kharin Yu. *Robustness in statistical pattern recognition*. Dordrecht: Kluwer Academic Publishers; 1996. 302 p.
2. Hampel FR, Ronchetti EM, Rousseeuw PJ, Stahel WA. *Robust statistics: the approach based on influence functions*. New York: Wiley; 1986. 350 p.
3. Kharin Yu. *Robustness in statistical forecasting*. New York: Springer; 2013. 356 p.
4. Maronna RA, Martin RD, Yohai VJ. *Robust statistics: theory and methods*. Chichester: Wiley; 2006. 420 p.
5. Rousseeuw PJ, Leroy AM. *Robust regression and other detection*. London: Wiley; 1987. 485 p.
6. Becker C, Fried R, Kuhnt S, editors. *Robustness and complex data structures*. New York: Springer; 2013. 283 p.
7. Fokianos K, Fried R, Kharin Yu, Voloshko V. Statistical analysis of multivariate discrete-valued time series. *Journal of Multivariate Analysis*. 2022;188:104805.
8. Kharin YuS, Voloshko VA. Robust estimation of AR coefficients under simultaneously influencing outliers and missing values. *Journal of Statistical Planning and Inference*. 2011;141:3276–3288. DOI: 10.1016/j.jspi.2011.04.015.
9. Kharin Yu, Voloshko V. Robust estimation for binomial conditionally non-linear autoregressive time series based on multivariate conditional frequencies. *Journal of Multivariate Analysis*. 2021;185:104777.
10. Kharin YuS, Petlitskii AI. A Markov chain of order s with r partial connections and statistical inference on its parameters. *Discrete Mathematics and Applications*. 2007;17(3):295–317.
11. Kharin Yu, Voloshko V. Statistical analysis of parsimonious high-order multivariate finite Markov chains based on sufficient statistics. *Journal of Multivariate Analysis*. 2025;208:105422. DOI: 10.1016/j.jmva.2025.105422.
12. Kharin YuS, Valoshka VA, Dernakova OV, Malugin VI, Kharin AY. Statistical forecasting of the dynamics of epidemiological indicators for COVID-19 incidence in the Republic of Belarus. *Journal of the Belarusian State University. Mathematics and Informatics*. 2020;3:36–50. Russian. DOI: 10.33581/2520-6508-2020-3-36-50.
13. Kim S, Park S, Na B, Yoon S. Spiking-YOLO: spiking neural network for energy-efficient object detection. arXiv:1903.06530 [Preprint]. 2019 [cited 2025 December 20]: [9 p.]. Available from: <https://arxiv.org/abs/1903.06530>.
14. Luo X, Yao M, Chou Y, Xu B, Li G. Integer-valued training and spike-driven inference spiking neural network for high-performance and energy-efficient object detection. arXiv:2407.20708 [Preprint]. 2024 [cited 2025 December 20]: [19 p.]. Available from: <https://arxiv.org/pdf/2407.20708>.
15. Niu W-H, Zhai R-B. A video human behavior recognition method based on improved 3D ResNet[J]. *Computer Engineering & Science*. 2023;45(10):1814–1821.
16. Lai HY, Hu C-C, Wen C-H, Wu J-X, Pai N-S, Yeh C-Y. Mel-scale frequency extraction and classification of dialect-speech signals with 1D CNN based classifier for gender and region recognition. *IEEE Access*. 2024;12:102962–102976. DOI: 10.1109/ACCESS.2024.3430296.
17. Haruna Y, Qin S, Chukkol AHA, Yusuf AA, Bello I, Lawan A. Exploring the synergies of hybrid convolutional neural network and vision transformer architectures for computer vision: a survey. *Engineering Applications of Artificial Intelligence*. 2025;144:110057. DOI: 10.1016/j.engappai.2025.110057.
18. Kang Hong H, Cho H. Cross-modal dynamic transfer learning for multimodal emotion recognition. *IEEE Access*. 2024;12:14324–14333. DOI: 10.1109/ACCESS.2024.3356185.

Received 29.01.2026 / revised 17.02.2026 / accepted 24.02.2026.

АСИМПТОТИЧЕСКАЯ НЕОТРИЦАТЕЛЬНАЯ КОРРЕЛИРОВАННОСТЬ РЯДА СТАТИСТИК, ИСПОЛЬЗУЕМЫХ ПРИ ПРОВЕРКЕ КАЧЕСТВА ГЕНЕРАТОРОВ СЛУЧАЙНЫХ ЧИСЕЛ

М. П. САВЕЛОВ¹⁾

¹⁾Математический институт им. В. А. Стеклова РАН,
ул. Губкина, 8, 119333, г. Москва, Россия

Аннотация. Рассматривается задача о проверке гипотезы, которая заключается в том, что тестируемая последовательность является последовательностью независимых случайных величин, имеющих известное полиномиальное распределение. Для решения этой задачи используются четыре типа статистик, представляющих собой обобщения статистик критериев пакетов NIST STS, TestU01 и др. Доказываются асимптотическая неотрицательная коррелированность многих из данных статистик, в том числе 10 статистик пакета NIST STS, и асимптотическая некоррелированность отдельных из них, в частности некоторых из 10 статистик пакета NIST STS.

Ключевые слова: NIST STS; асимптотически неотрицательно коррелированные статистики; асимптотически некоррелированные статистики; совместное распределение статистик; предельные распределения; пакет TestU01; генераторы случайных чисел.

Благодарность. Автор выражает признательность доктору физико-математических наук А. М. Зубкову за постоянную помощь в подготовке работы.

ASYMPTOTIC NON-NEGATIVE CORRELATION OF SOME STATISTICS USED IN TESTING THE QUALITY OF RANDOM NUMBER GENERATORS

M. P. SAVELOV^a

^aSteklov Mathematical Institute of Russian Academy of Sciences,
8 Gubkina Street, Moscow 119333, Russia

Abstract. The problem of testing the hypothesis, stating that the tested sequence consists of independent random variables with a given polynomial distribution, is considered. To solve this problem four types of statistics that are generalisations of statistics of tests of the NIST STS, TestU01 and other packages are used. It is proved that many of these statistics, including 10 statistics of the NIST STS, are asymptotically non-negatively correlated, and some of them, in particular some of the 10 statistics of the NIST STS, are asymptotically uncorrelated.

Keywords: NIST STS; asymptotically non-negatively correlated statistics; asymptotically uncorrelated statistics; joint distribution of statistics; limit distributions; TestU01 package; random number generators.

Acknowledgements. The author expresses his gratitude to doctor of science (physics and mathematics) A. M. Zubkov for constant assistance in preparing work.

Образец цитирования:

Савелов МП. Асимптотическая неотрицательная коррелированность ряда статистик, используемых при проверке качества генераторов случайных чисел. Журнал Белорусского государственного университета. Математика. Информатика. 2026;1:16–21 (на англ.).
EDN: AJDZWV

For citation:

Savelov MP. Asymptotic non-negative correlation of some statistics used in testing the quality of random number generators. Journal of the Belarusian State University. Mathematics and Informatics. 2026;1:16–21.
EDN: AJDZWV

Автор:

Максим Павлович Савелов – кандидат физико-математических наук; старший научный сотрудник отдела дискретной математики.

Author:

Maxim P. Savelov, PhD (physics and mathematics); senior researcher at the department of discrete mathematics.
savelovmp@gmail.com
<https://orcid.org/0000-0001-7751-6281>

Introduction

One of the most famous tools used to test random and pseudorandom number generators is the National Institute of Standards and Technology statistical test suite (here and further – NIST STS) [1]. In recent years many papers have been published on this topic (e. g., articles [2–12]).

In this study we consider four types of statistics defined in works [6; 7]. They are generalisations of statistics of some tests of the NIST STS and other packages. We analyse the question about asymptotic uncorrelation and asymptotic non-negative correlation of these statistics.

Main results

We need some definitions from papers [6; 7]. Let us formulate them.

Let $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ be random variables, taking values in the set $\{0, 1, \dots, R-1\}$. Consider the hypothesis H_0 that the tested sequence consists of independent random variables with a known polynomial distribution. Let N_{lb} and L_{sb} be natural numbers (they are the parameters, by which the statistics T_{lb} and T_{sb} will be constructed). We analyse the case when N_{lb} and L_{sb} are fixed and n tends to infinity. Put

$$L_{lb} = \left\lfloor \frac{n}{N_{lb}} \right\rfloor, \quad N_{sb} = \left\lfloor \frac{n}{L_{sb}} \right\rfloor.$$

If the last $n - N_{lb} L_{lb}$ elements are discarded from the sequence $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$, then the remaining elements may be divided into N_{lb} «long» non-intersecting blocks of length L_{lb} : the first block consists of random variables $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{L_{lb}}$, the second block consists of random variables $\varepsilon_{L_{lb}+1}, \varepsilon_{L_{lb}+2}, \dots, \varepsilon_{2L_{lb}}$, etc. Similarly, we can discard the last $n - L_{sb} N_{sb}$ elements and split the remaining ones into N_{sb} «short» blocks of length L_{sb} .

Next we need to fix natural numbers $m_{\text{sum}}, m_{lb}, K_{sb}$ and also three functions $f_{\text{sum}} : \{0, 1, \dots, R-1\}^{m_{\text{sum}}} \rightarrow \mathbb{R}$, $f_{lb} : \{0, 1, \dots, R-1\}^{m_{lb}} \rightarrow \mathbb{R}$ and $f_{sb} : \{0, 1, \dots, R-1\}^{L_{sb}} \rightarrow \mathbb{R}$. Let us partition the set of values of the function f_{sb} into $K_{sb} + 1$ non-empty disjoint subsets: $f_{sb}(\{0, 1, \dots, R-1\}^{L_{sb}}) = \coprod_{i=0}^{K_{sb}} \alpha_{sb}(i)$. We will be interested in the case, when the numbers $R, p_0, \dots, p_{R-1}, m_{\text{sum}}, m_{lb}, N_{lb}, L_{sb}, K_{sb}$, the functions $f_{\text{sum}}, f_{lb}, f_{sb}$ and the sets $\alpha_{sb}(0), \dots, \alpha_{sb}(K_{sb})$ are fixed, do not depend on n and n tends to infinity.

Put

$$E_{\text{sum}} = \mathbf{E} f_{\text{sum}}(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{m_{\text{sum}}}),$$

$$\sigma_{\text{sum}}^2 = \mathbf{D} f_{\text{sum}}(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{m_{\text{sum}}}) + 2 \sum_{i=2}^{m_{\text{sum}}} \text{cov}(f_{\text{sum}}(\varepsilon_i, \varepsilon_{i+1}, \dots, \varepsilon_{m_{\text{sum}}-1}), f_{\text{sum}}(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{m_{\text{sum}}-i})),$$

where it's supposed that $\sigma_{\text{sum}} \geq 0$.

Definition 1. If $\sigma_{\text{sum}} > 0$, then a statistic of the form

$$T_{\text{sum}} = \frac{\sum_{i=1}^{n - m_{\text{sum}} + 1} (f_{\text{sum}}(\varepsilon_i, \varepsilon_{i+1}, \dots, \varepsilon_{i+m_{\text{sum}}-1}) - E_{\text{sum}})}{\sigma_{\text{sum}} \sqrt{n - m_{\text{sum}} + 1}} \quad (1)$$

is called a summing statistic.

Note that in paper [5], under the assumption that the tested sequence is binary, statistics similar to summing statistics are considered.

Put

$$W_k = \sum_{j=L_{lb}(k-1)+1}^{L_{lb}k - m_{lb} + 1} f_{lb}(\varepsilon_j, \varepsilon_{j+1}, \dots, \varepsilon_{j+m_{lb}-1}), \quad 1 \leq k \leq N_{lb},$$

$$E_{lb} = \mathbf{E} f_{lb}(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{m_{lb}}),$$

$$\sigma_{lb}^2 = \mathbf{D} f_{lb}(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{m_{lb}}) + 2 \sum_{i=2}^{m_{lb}} \text{cov}(f_{lb}(\varepsilon_i, \varepsilon_{i+1}, \dots, \varepsilon_{i+m_{lb}-1}), f_{lb}(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{m_{lb}-i})),$$

where it is supposed that $\sigma_{lb} \geq 0$. Note that $\mathbf{E} W_k = (L_{lb} - m_{lb} + 1) E_{lb}$, $1 \leq k \leq N_{lb}$.

Definition 2. If $\sigma_{lb} > 0$, then a statistic of the form

$$T_{lb} = \frac{\sum_{k=1}^{N_{lb}} (W_k - (L_{lb} - m_{lb} + 1) E_{lb})^2}{L_{lb} \sigma_{lb}^2}$$

is called a long-block statistic with N_{lb} blocks.

For $0 \leq j \leq K_{sb}$ put

$$w(j) = \sum_{i=1}^{N_{sb}} I_{f_{sb}}(\varepsilon_{L_{sb}(i-1)+1}, \varepsilon_{L_{sb}(i-1)+2}, \dots, \varepsilon_{L_{sb}i}) \in \alpha_{sb}(j),$$

$$E_{sb}(j) = \mathbf{P}(f_{sb}(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{L_{sb}}) \in \alpha_{sb}(j)).$$

Note that $\mathbf{E}w(j) = N_{sb} E_{sb}(j)$, $0 \leq j \leq K_{sb}$.

Definition 3. If $E_{sb}(j) > 0$ for all $0 \leq j \leq K_{sb}$, then a statistic of the form

$$T_{sb} = \sum_{j=0}^{K_{sb}} \frac{(w(j) - N_{sb} E_{sb}(j))^2}{N_{sb} E_{sb}(j)}$$

is called a short-block statistic with blocks of length L_{sb} .

Next for each $1 \leq q \leq Q_{\text{quad}}$ fix a natural number $m_{\text{sum}}^{[q]}$ and a function $f_{\text{sum}}^{[q]} : \{0, 1, \dots, R-1\}^{m_{\text{sum}}^{[q]}} \rightarrow \mathbb{R}$ and put

$$T_{\text{sum}}^{[q]} = \frac{\sum_{i=1}^{n - m_{\text{sum}}^{[q]} + 1} \left(f_{\text{sum}}^{[q]}(\varepsilon_i, \varepsilon_{i+1}, \dots, \varepsilon_{i+m_{\text{sum}}^{[q]}-1}) - E_{\text{sum}}^{[q]} \right)^2}{\sigma_{\text{sum}}^{[q]} \sqrt{n - m_{\text{sum}}^{[q]} + 1}}, \quad (2)$$

where the quantities $E_{\text{sum}}^{[q]}$ and $\sigma_{\text{sum}}^{[q]}$ are defined in the same way as the quantities E_{sum} and σ_{sum} in summing statistic (1). The statistics $T_{\text{sum}}^{[1]}, \dots, T_{\text{sum}}^{[Q_{\text{quad}}]}$ are summing statistics.

Definition 4. Consider a positive integer τ_{quad} and real numbers $d_{\text{quad}}(i, q)$, $1 \leq i \leq \tau_{\text{quad}}$, $1 \leq q \leq Q_{\text{quad}}$. A statistic of the form

$$T_{\text{quad}} = \sum_{i=1}^{\tau_{\text{quad}}} \left(\sum_{q=1}^{Q_{\text{quad}}} d_{\text{quad}}(i, q) T_{\text{sum}}^{[q]} \right)^2$$

is called a quadratic statistic constructed from the statistics $T_{\text{sum}}^{[1]}, \dots, T_{\text{sum}}^{[Q_{\text{quad}}]}$.

Remark. The relationship between the summing, long-block, short-block and quadratic statistics and the statistics from NIST STS [1], TestU01 and other packages was discussed in papers [6; 7]. In particular, the following was proved with respect to the NIST STS in article [6]. Let $R = 2$, $p_0 = p_1 = \frac{1}{2}$ and the hypothesis H_0 be true. The statistics T_{fr} , T_{templ} of tests «Frequency test within a block» and «Non-overlapping template matching test» are long-block statistics, the statistics T_{longrun} , T_{matrix} , T_{lincompl} of tests «Test for the longest run of ones in a block», «Binary matrix rank test», «Linear complexity test» are short-block statistics (taking into account the caveat from example 7 [6]), the statistic T_{mon} of test «Monobit test» is a summing one, the statistic T_{runs} of test «Runs test» coincides up to $o_p(1)$ with the summing statistic, statistics T_{serial1} , T_{serial2} of the test «Serial test» and the statistic T_{entr} of test «Approximate entropy test» coincide up to $o_p(1)$ with some quadratic statistics.

We will call the numbers m_{sum} , m_{lb} , N_{lb} , L_{sb} , K_{sb} , the functions f_{sum} , f_{lb} , f_{sb} and the sets $\alpha_{sb}(0), \dots, \alpha_{sb}(K_{sb})$ the parameters of the statistics T_{sum} , T_{lb} , T_{sb} . The numbers Q_{quad} , τ_{quad} and $d_{\text{quad}}(i, q)$, $1 \leq i \leq \tau_{\text{quad}}$, $1 \leq q \leq Q_{\text{quad}}$, as well as the parameters of the statistics $T_{\text{sum}}^{[1]}, \dots, T_{\text{sum}}^{[Q_{\text{quad}}]}$, are called the parameters of the statistic T_{quad} . Further we will assume that they do not depend on n and are fixed.

Let us consider Q triples of statistics $(T_{\text{sum}}^{[q]}, T_{lb}^{[q]}, T_{sb}^{[q]})$, $1 \leq q \leq Q$, where $T_{\text{sum}}^{[q]}$, $T_{lb}^{[q]}$ and $T_{sb}^{[q]}$ are the summing, long-block and short-block statistics, respectively. We will use the following notation. As in equality (2), all

parameters related to $T_{\text{sum}}^{[q]}$ will be marked with the subscript «sum» and the superscript $[q]$. Similarly, $N_{lb}^{[q]}$ is the number of long blocks used to construct the statistic $T_{lb}^{[q]}$, $L_{sb}^{[q]}$ is the length of short blocks used in constructing the statistic $T_{sb}^{[q]}$, etc. The triple of statistics $(T_{\text{sum}}^{[q]}, T_{lb}^{[q]}, T_{sb}^{[q]})$ is constructed from functions $f_{\text{sum}}^{[q]}, f_{lb}^{[q]}, f_{sb}^{[q]}$.

Consider J quadratic statistics $T_{\text{quad}}^{[1]}, \dots, T_{\text{quad}}^{[J]}$ constructed from $T_{\text{sum}}^{[1]}, \dots, T_{\text{sum}}^{[Q]}$:

$$T_{\text{quad}}^{[j]} = \sum_{i=1}^{\tau_{\text{quad}}^{[j]}} \left(\sum_{q=1}^Q d_{\text{quad}}^{[j]}(i, q) T_{\text{sum}}^{[q]} \right)^2.$$

Definition 5. Let $\vec{T}$ and $\vec{T}^*$ be multivariate statistics constructed from the sample $\varepsilon_1, \dots, \varepsilon_n$. Following the paper [9], we say that they are asymptotically uncorrelated if there exist random vectors $\vec{\zeta}$ and $\vec{\zeta}^*$ such that $(\vec{T}, \vec{T}^*) \xrightarrow{d} (\vec{\zeta}, \vec{\zeta}^*)$ as $n \rightarrow \infty$ and $\text{cov}(\vec{\zeta}, \vec{\zeta}^*) = 0$. We say that univariate statistics T and T^* are asymptotically non-negatively correlated if there exist random variables ζ and ζ^* such that $(T, T^*) \xrightarrow{d} (\zeta, \zeta^*)$ as $n \rightarrow \infty$ and the inequality $\text{cov}(\zeta, \zeta^*) \geq 0$ is satisfied.

Theorem. Let the hypothesis H_0 be true. If the parameters of the statistics $T_{\text{sum}}^{[1]}, T_{lb}^{[1]}, T_{sb}^{[1]}, \dots, T_{\text{sum}}^{[Q]}, T_{lb}^{[Q]}, T_{sb}^{[Q]}, T_{\text{quad}}^{[1]}, \dots, T_{\text{quad}}^{[J]}$ are fixed, then the vectors $(T_{\text{sum}}^{[1]}, \dots, T_{\text{sum}}^{[Q]})$ and $(T_{lb}^{[1]}, T_{sb}^{[1]}, \dots, T_{lb}^{[Q]}, T_{sb}^{[Q]}, T_{\text{quad}}^{[1]}, \dots, T_{\text{quad}}^{[J]})$ are asymptotically uncorrelated, and the components of the second of these vectors are pairwise asymptotically non-negatively correlated.

Note that using theorem 1 [7] one can find the covariance matrix of the vector to which the vector $(T_{\text{sum}}^{[1]}, T_{lb}^{[1]}, T_{sb}^{[1]}, \dots, T_{\text{sum}}^{[Q]}, T_{lb}^{[Q]}, T_{sb}^{[Q]}, T_{\text{quad}}^{[1]}, \dots, T_{\text{quad}}^{[J]})$ converges in distribution, not only in the case of H_0 but also in the case of alternatives H_1 «close» to H_0 (see examples of such alternatives in article [6]).

Further explicit formulas for the statistics $T_{\text{mon}}, T_{\text{runs}}, T_{\text{fr}}, T_{\text{longrun}}, T_{\text{matrix}}, T_{\text{templ}}, T_{\text{lincompl}}, T_{\text{serial1}}, T_{\text{serial2}}, T_{\text{entr}}$ of the NIST STS, as well as the definition of the parameters of these statistics, are presented in papers [6; 7]. In particular, N_{fr} is the number of blocks in the long-block statistic T_{fr} , L_{matrix} is the length of blocks of short-block statistic of the test «Matrix test», m_{templ} is the length of the template from which the long-block statistic T_{templ} is constructed, etc.

Corollary. Let $R = 2$, $p_0 = p_1 = \frac{1}{2}$ and the hypothesis H_0 be true. If the parameters of the statistics $T_{\text{mon}}, T_{\text{runs}}, T_{\text{fr}}, T_{\text{longrun}}, T_{\text{matrix}}, T_{\text{templ}}, T_{\text{lincompl}}, T_{\text{serial1}}, T_{\text{serial2}}, T_{\text{entr}}$ are fixed, then the components of the vector $(T_{\text{mon}}, T_{\text{runs}}, T_{\text{fr}}, T_{\text{longrun}}, T_{\text{matrix}}, T_{\text{templ}}, T_{\text{lincompl}}, T_{\text{serial1}}, T_{\text{serial2}}, T_{\text{entr}})$ are pairwise asymptotically non-negatively correlated, the statistics T_{mon} and T_{runs} and the vector $(T_{\text{fr}}, T_{\text{longrun}}, T_{\text{matrix}}, T_{\text{templ}}, T_{\text{lincompl}}, T_{\text{serial1}}, T_{\text{serial2}}, T_{\text{entr}})$ are pairwise asymptotically uncorrelated.

The assertion of corollary intersects with the results of works [9–12] and with the results of some other papers of M. P. Savelov and complements them.

Lemma 1. Suppose that conditions of corollary hold and the parameter m_{serial} of the statistic T_{serial2} satisfies the inequality $m_{\text{serial}} \geq 2$. Then statistics T_{fr} and T_{serial2} are asymptotically uncorrelated.

Let us prove the results formulated above. We will need the following statement, which is proven in article [9, formula (15)].

Lemma 2. Let (ξ, η) be a two-dimensional random vector with a normal distribution and zero mean. Then

$$\text{cov}(\xi, \eta^2) = 0, \text{cov}(\xi^2, \eta^2) = 2\text{cov}(\xi, \eta), \text{cov}(\xi^2, \xi) = 0.$$

Proof of the theorem. Put

$$N = \prod_{q=1}^Q N_{lb}^{[q]}, h = \prod_{q=1}^Q L_{sb}^{[q]}, s = h + \max_{1 \leq q \leq Q} \max(m_{\text{sum}}^{[q]}, m_{lb}^{[q]}) - 1.$$

As noted in article [7], the hypothesis H_0 satisfies the conditions $\mathbf{I}(N, s, h)$ and $\mathbf{II}(N, s, h, \vec{0})$ defined in work [7]. Let the vectors $\zeta_1^{[q]}, \dots, \zeta_{N_{lb}^{[q]}}^{[q]}$, $1 \leq q \leq Q$, be the same as in paper [7] (we do not give the explicit form of these vectors due to its cumbersomeness). Applying theorem 1 of article [7], we obtain that the vector

$$\left(T_{\text{sum}}^{[1]}, T_{lb}^{[1]}, T_{sb}^{[1]}, \dots, T_{\text{sum}}^{[Q]}, T_{lb}^{[Q]}, T_{sb}^{[Q]}, T_{\text{quad}}^{[1]}, \dots, T_{\text{quad}}^{[J]} \right)$$

converges in distribution to the vector

$$\left(\sum_{k=1}^{N_b^{[1]}} \zeta_k^{[1]} (K_{sb}^{[1]} + 2), N_{lb}^{[1]} \sum_{k=1}^{N_b^{[1]}} \left(\zeta_k^{[1]} (K_{sb}^{[1]} + 1) \right)^2, L_{sb}^{[1]} \sum_{j=0}^{K_{sb}^{[1]}} \left(\sum_{k=1}^{N_b^{[1]}} \zeta_k^{[1]}(j) \right)^2, \dots, \right. \\ \left. \sum_{k=1}^{N_b^{[Q]}} \zeta_k^{[Q]} (K_{sb}^{[Q]} + 2), N_{lb}^{[Q]} \sum_{k=1}^{N_b^{[Q]}} \left(\zeta_k^{[Q]} (K_{sb}^{[Q]} + 1) \right)^2, L_{sb}^{[Q]} \sum_{j=0}^{K_{sb}^{[Q]}} \left(\sum_{k=1}^{N_b^{[Q]}} \zeta_k^{[Q]}(j) \right)^2, \theta^{[1]}, \dots, \theta^{[J]} \right)$$

as $n \rightarrow \infty$, where

$$\theta^{[j]} = \sum_{i=1}^{\tau_{\text{quad}}^{[j]}} \left(\sum_{q=1}^Q d_{\text{quad}}^{[j]}(i, q) \sum_{k=1}^{N_b^{[q]}} \zeta_k^{[q]} (K_{sb}^{[q]} + 2) \right)^2, \quad 1 \leq j \leq J,$$

and, in addition, the vectors $\zeta_1^{[q]}, \dots, \zeta_{N_b^{[q]}}^{[q]}, 1 \leq q \leq Q$, are linear functions of the centered Gaussian random vector ξ defined in paper [7]. From here, by virtue of lemma 2, the statement of theorem follows.

Proof of the corollary. By remark, there exist statistics $T_{\text{sum}}^{[1]}, T_{\text{sum}}^{[2]}, T_{lb}^{[1]}, T_{lb}^{[2]}, T_{sb}^{[1]}, T_{sb}^{[2]}, T_{sb}^{[3]}, T_{\text{quad}}^{[1]}, T_{\text{quad}}^{[2]}, T_{\text{quad}}^{[3]}$ such that

$$T_{\text{mon}} = T_{\text{sum}}^{[1]}, T_{\text{runs}} = T_{\text{sum}}^{[2]} + o_P(1), T_{\text{fr}} = T_{lb}^{[1]}, T_{\text{templ}} = T_{lb}^{[2]}, \\ T_{\text{longrun}} = T_{sb}^{[1]}, T_{\text{matrix}} = T_{sb}^{[2]}, T_{\text{lincompl}} = T_{sb}^{[3]}, T_{\text{serial1}} = T_{\text{quad}}^{[1]} + o_P(1), \\ T_{\text{serial2}} = T_{\text{quad}}^{[2]} + o_P(1), T_{\text{entr}} = T_{\text{quad}}^{[3]} + o_P(1).$$

Put $Q = 3, J = 3, T_{\text{sum}}^{[3]} = T_{\text{sum}}^{[1]}, T_{lb}^{[3]} = T_{lb}^{[1]}$. Applying theorem, we obtain that the vectors $(T_{\text{mon}}, T_{\text{runs}})$ and $(T_{\text{fr}}, T_{\text{longrun}}, T_{\text{matrix}}, T_{\text{templ}}, T_{\text{lincompl}}, T_{\text{serial1}}, T_{\text{serial2}}, T_{\text{entr}})$ are asymptotically uncorrelated, and the components of the second of these vectors are pairwise asymptotically non-negatively correlated. It remains to show that T_{mon} and T_{runs} are asymptotically uncorrelated.

Consider the statistic T_{templ}^* constructed from $N_{\text{templ}}^* = N_{\text{fr}}$ long blocks and some template of length $m_{\text{templ}}^* \geq 2$. Applying assertion 1 of article [9] to the vector $(T_{\text{mon}}, T_{\text{fr}}, T_{\text{templ}}^*, T_{\text{runs}})$, we obtain that this vector converges (as $n \rightarrow \infty$) in distribution to a vector, first and last components of which are uncorrelated. Note that in [9] it is additionally assumed that n is divisible by N_{fr} [9, p. 55] but an analysis of the proof shows that this condition can be discarded. Thus, the statistics T_{mon} and T_{runs} are asymptotically uncorrelated.

Corollary is proved.

Proof of lemma 1. Reasoning in the same way as above in the proof of theorem and applying lemma 3 of work [7], we obtain that the statistics T_{fr} and T_{serial2} are asymptotically independent and, as a consequence, asymptotically uncorrelated (see a similar result in lemma 2 [12]).

Lemma 1 is proved.

Conclusions

We considered four types of statistics that are generalisations of statistics of tests of the NIST STS, TestU01 and other packages. We proved that many of these statistics are asymptotically non-negatively correlated and some of them are asymptotically uncorrelated. We obtained asymptotic non-negative correlation of 10 statistics of the NIST STS and asymptotic uncorrelation of some of these 10 statistics.

References

1. Rukhin AL, Soto J, Nechvatal JR, Smid ME, Barker EB, Leigh SD, et al. *A statistical test suite for random number generators and pseudorandom number generators for cryptographic applications*. Gaithersburg: National Institute of Standards and Technology; 2010. 131 p.
2. Voloshko VA, Kharin YuS, Trubey AI. On power comparison for some tests on pure randomness under Markov high-order dependencies. In: Kharin YuS, editor. *Computer data analysis and modeling: stochastics and data science. Proceedings of the 13th International conference; 2022 September 6–10; Minsk, Belarus*. Minsk: Belarusian State University; 2022. p. 211–217.

3. Zubkov AM, Serov AA. Experimental study of NIST statistical test suite ability to detect long repetitions in binary sequences. *Matematicheskie voprosy kriptografii*. 2023;14(2):137–145. DOI: 10.4213/mvk443.
4. Kharin YuS, Zubkov AM. On statistical testing of composite hypotheses on s -dimensional uniform probability distribution of binary sequences. *Diskretnaya matematika*. 2024;36(1):116–135. Russian. DOI: 10.4213/dm1806.
5. Voloshko VA. On the asymptotic properties of the family of χ^2 -tests of pure randomness of a binary sequence. In: Kharin YuS. *Theoretical and applied cryptography. Proceedings of the 2nd International scientific conference; 2023 October 19–20; Minsk, Belarus*. Minsk: Belarusian State University; 2023. p. 15–43. Russian.
6. Savelov MP. The limit joint distributions of statistics of tests of the NIST package and their generalisations. *Diskretnaya matematika*. 2024;36(2):71–116. Russian. DOI: 10.4213/dm1824.
7. Savelov MP. Asymptotic independence of statistics of tests of the NIST package and their generalisations. *Diskretnaya matematika*. 2025;37(1):76–111. Russian. DOI: 10.4213/dm1862.
8. Savelov MP. Analogue of the Berry – Esseen inequality for the joint cumulative distribution function of several statistics. *Diskretnaya matematika*. 2025;37(2):66–108. Russian. DOI: 10.4213/dm1875.
9. Savelov MP. The limit joint distributions of statistics of four tests of the NIST package. *Discrete Mathematics and Applications*. 2023;33(1):55–64. DOI: 10.1515/dma-2023-0006.
10. Savelov MP. Limit joint distribution of the statistics of «Monobit test», «Frequency test within a block» and «Test for the longest run of ones in a block». *Discrete Mathematics and Applications*. 2024;34(5):291–301. DOI: 10.1515/dma-2024-0026.
11. Savelov MP. Limit joint distribution of the statistics of «Monobit test», «Frequency test within a block» and «Binary matrix rank test». *Discrete Mathematics and Applications*. 2025;35(4):249–259. DOI: 10.1515/dma-2025-0019.
12. Savelov MP. Limit joint distribution of «Monobit test», «Frequency test within a block», and «Serial test» statistics. *Diskretnaya matematika*. 2023;35(1):88–106. Russian. DOI: 10.4213/dm1744.

Received 24.12.2025 / revised 19.01.2026 / accepted 29.01.2026.

НАДЕЖНОСТЬ ДВУХУРОВНЕВОГО ПОДХОДА К ТЕСТИРОВАНИЮ В НАБОРЕ ТЕСТОВ NIST SP 800-22

А. А. СЕРОВ¹⁾

¹⁾Математический институт им. В. А. Стеклова РАН, ул. Губкина, 8, 119333, г. Москва, Россия

Аннотация. Рассматривается предложенный в наборе тестов NIST SP 800-22 двухуровневый подход к тестированию генераторов случайных чисел, который состоит в подсчете количества последовательностей, прошедших основной тест, и проверке распределения p -значений с помощью хи-квадрат-теста. Данный подход может повысить надежность статистического теста, однако он чувствителен к погрешностям аппроксимации, возникающим при вычислении p -значений. Показано, что для последовательностей, порождаемых с помощью симметричного алгоритма блочного шифрования (AES), двухуровневый подход к тестированию не является надежным. Систематическая погрешность при вычислении p -значений зависит только от погрешности аппроксимации точного распределения статистики эквивалентным ему теоретическим распределением и от числа бит n в тестируемых последовательностях. Для надежного двухуровневого тестирования такая погрешность не должна превосходить

$\frac{\sigma}{N} = \frac{1}{k} \sqrt{\frac{k-1}{N}}$, где $\sigma = \sqrt{\frac{1}{k} \left(1 - \frac{1}{k}\right)} N$ – стандартное отклонение от среднего числа частиц в ячейке в равновероятной

схеме размещения N частиц по k ячейкам. Подобные эвристические соображения и проведенные эксперименты позволяют предположить, что, например, при $n = 2^{20}$ в двухуровневом частотном тесте (*Frequency test*) из набора тестов NIST SP 800-22 количество тестируемых последовательностей N не должно превышать 26 000. Для полного исключения систематической погрешности, возникающей при проведении частотного теста и определении номера ячейки из k возможных ячеек (непересекающихся подынтервалов отрезка $[0, 1]$), которому принадлежит p -значение, предложены двусторонние оценки квантилей биномиального закона.

Ключевые слова: случайная последовательность; псевдослучайная последовательность; статистическое тестирование; надежность статистического теста; биномиальное распределение; двусторонняя оценка.

RELIABILITY OF TWO-LEVEL TESTING APPROACH IN THE NIST SP 800-22 TEST SUITE

A. A. SEROV^a

^a*Steklov Mathematical Institute, Russian Academy of Sciences, 8 Gubkina Street, Moscow 119333, Russia*

Abstract. The two-level approach for testing random number generators involving the well-known NIST SP 800-22 test suite, i. e. counting the sequences passing a basic test and checking the p -values distribution with a chi-square test, was considered. Such an approach may increase the reliability of the test. However, it is sensitive to the approximation error introduced by the computing of p -values. In this paper it is shown that for AES-based sequences two-level testing approach is not reliable too. Systematic error in the computing of the p -values is dependent only on the ac-

Образец цитирования:

Серов АА. Надежность двухуровневого подхода к тестированию в наборе тестов NIST SP 800-22. *Журнал Белорусского государственного университета. Математика. Информатика.* 2026;1:22–35 (на англ.).
EDN: BPTBQY

For citation:

Serov AA. Reliability of two-level testing approach in the NIST SP 800-22 test suite. *Journal of the Belarusian State University. Mathematics and Informatics.* 2026;1:22–35.
EDN: BPTBQY

Автор:

Александр Александрович Серов – кандидат физико-математических наук; научный сотрудник отдела дискретной математики.

Author:

Aleksandr A. Serov, PhD (physics and mathematics); researcher at the department of discrete mathematics.
serov@mi-ras.ru
<https://orcid.org/0000-0001-7465-1701>

curacy of approximation of the exact distribution of statistic by its theoretical counterpart and the number of bits in the analysed sequences n . For a reliable second-level test, this error should be smaller, or at least, approximately equal to

$$\frac{\sigma}{N} = \frac{1}{k} \sqrt{\frac{k-1}{N}}, \text{ where } \sigma = \sqrt{\frac{1}{k} \left(1 - \frac{1}{k}\right)} N \text{ is the standard deviation from the mean number of particles in a bin for equiprobable scheme of allocation } N \text{ particles in } k \text{ bins.}$$

Such heuristic assumptions and carried out experiments suggest that for example in the second-level test of the Frequency test of NIST SP 800-22 test suite with $n = 2^{20}$ the number of tested sequences N should not exceed 26 000. To completely eliminate the systematic error appearing in the Frequency test when determining the number of bin from k bins (disjoint subintervals of the interval $[0, 1]$) to which the p -value belongs the two-sided estimates of the quantiles of the binomial law are proposed.

Keywords: random sequence; pseudorandom sequence; statistical testing; reliability of statistical test; binomial distribution; two-sided estimate.

Introduction

Random sequences are used in a large variety of areas, such as quantum mechanics, game theory, statistics, cryptography and so on. These sequences may be generated either by physical sources or deterministic algorithms. Random number generators (RNGs) represent a fundamental component in many applications, they are essential for cryptographic systems (see, for example, [1]). The security of many cryptographic schemes and protocols is based on the perfect randomness of RNG outputs. RNGs are classified into two types – pseudo (deterministic) and true (non-deterministic) RNGs (further – PRNGs and TRNGs, respectively). In general, TRNGs based on some random physical phenomena, may be used directly as random bit sources or generate seeds for PRNGs, and PRNGs extend the seeds to produce deterministic long sequences.

For any type of RNG, statistical hypothesis tests have been widely employed to assess the quality of the RNG, which evaluate whether the output sequences conform with the given null hypothesis (e. g., the elements of the sequence independent and uniformly distributed) or not. In addition, statistical randomness tests are also used to evaluate the outputs of other cryptographic primitives such as block ciphers and hash functions, to preliminarily validate the indistinguishability of their outputs from a equiprobable random mapping. The quality checking of binary sequences usually is based on some well-known batteries of tests, each of which is composed of a serial of tests, include Diehard¹ proposed by G. Marsaglia, SP 800-22 [2] standardised by National Institute of Standards and Technology (further – NIST SP 800-22) or a software library TestU01 [3].

Note that a test for randomness may be interpreted only in a probabilistic way. Looking at a sequence of all zeros, we say that the sequence is not random at all, though for an ideal RNG this sequence has the same probability of any other sequence of the same length; on the contrary, a RNG that is not able to generate this sequence is not ideal.

The main problem of the statistical testing is the interpretation of the results. Roughly speaking, while a failed test is a serious indicator for the weakness of a RNG, a passed test does not provide a direct positive proof for the quality of a RNG. From a mathematical point of view a test may be considered as a function of a sequence of n elements (e. g., a sequence of n bits) with output value in the interval $[0, 1]$, called a p -value. In null-hypothesis significance testing, the p -value is the probability of obtaining test results at least as extreme as the result actually observed, under the assumption that the null hypothesis is correct.

In general case, if the null hypothesis is true, the p -value based on a continuous test statistic has a uniform distribution over the interval $[0, 1]$, regardless of the sample size of the experiment. In contrast, the distribution of the p -value under the alternative hypothesis is a function of both sample size and the true value or range of true values of the tested parameter of the true distribution. In the discrete case, then the test statistic distribution is not continuous and the null hypothesis is true, i. e. all elements of the input sequence are independent and drawn according to the uniform distribution, the cumulative distribution function (CDF) $F_p(x)$ of p -value coincides with the value of its argument in a discrete set of points from the interval $[0, 1]$. If the elements of the generated sequence is non-independent or not distributed according to the uniform distribution, then the CDF of p -value is not known as a rule.

NIST SP 800-22 test suite

In this paper the most commonly used statistical test suite, the NIST SP 800-22 test suite [2], is considered. It is well known that the NIST statistical test suite was used for the evaluation of AES candidate algorithms. This

¹Marsaglia G. Diehard: a battery of tests of randomness [Electronic resource]. URL: <http://stat.fsu.edu/geo/diehard.html> (date of access: 21.06.2025).

statistical test suite is build for analysing the randomness properties of sequences and generators, is composed of 15 tests, and provides comprehensive evaluation for different randomness aspects of assessed sequences. All NIST SP 800-22 tests are listed in table 1.

Table 1

All NIST SP 800-22 tests with brief descriptions

Number	Test name	Used statistics
1	Frequency test	Normalised modulus of the difference between frequencies of 1 and 0
2	Block frequency test	Chi-square statistic of 1 frequencies in adjacent non-intersecting 128-bit blocks
3	Cumulative sums test	Maximal deviation from 0 for partial sums of ± 1 walk constructed by the binary sequence
4	Runs test	The total number of 1-runs and 0-runs
5	Longest run test	Pearson statistic of maximal lengths of 1-runs frequencies in adjacent non-intersecting 10^4 -bit blocks
6	Binary matrix rank test	Pearson statistic of binary 32×32 -matrices ranks of frequencies formed from adjacent non-intersecting 1024-bit blocks of the sequence
7	Discrete Fourier transform test	Normalised difference between the number of discrete Fourier transform coefficients of n -bit sequence, exceeding $\sqrt{n \ln 20}$ in absolute value, and $\frac{0.95n}{2}$
8	Overlapping template matching test	Pearson statistic of 9-bit 1-series frequencies in 1032-bit adjacent blocks with overlappings
9	Maurer's universal statistical test	Sum of base 2 logarithms of distances between equal non-intersecting 7-bit blocks
10	Approximate entropy test	Difference of logarithmic frequencies statistics of 10- and 11-bit segments with overlappings, the reference distribution is the χ^2 distribution
11	Serial test	Differences in Pearson statistics of frequencies of all 14-, 15- and 16-bit segments with overlappings
12	Linear complexity test	Pearson statistic of shortest LSR lengths frequencies generating 500-bit segments of the sequence
13	Non-overlapping template matching test	Chi-square statistics of aperiodic fixed segments frequencies in 8 adjacent non-overlapping blocks
14	Random excursions test	Pearson statistics of random walk cycles frequencies with fixed numbers visiting of $-4, -3, \dots, 4$ states
15	Random excursions variant test	The total number of times that the given 18 states from -9 to 9 is visited during the entire random walk

The main reason for using the NIST SP 800-22 test suite, from an engineering point of view, is that it has several appealing properties. First, the suite is composed of several different tests, each of them is applied to the same input sequence of n bits (for many of the tests the assumption has been made that the sequence length n is a value from 10^3 to 10^7), searching for a specific statistical feature and expressing it as numerical quantity of p -value². Second, the suite is composed by a number of well-known tests and, for all of them, an exhaustive mathematical treatment is available. The source code of all the tests in the suite is public available.

The purpose of this paper is to consider the testing strategy proposed in section 4 of the NIST publication [2] and discuss under which assumptions this strategy increases the reliability and when, on the other hand, produces incorrect results, i. e. the empirical significance level α does not correspond to the theoretical one. In that context, a reliable test should be understood as a test such that the probability of a false positive (type I error) is agreed with the expected one.

In the beginning of the testing process, the whole bit sequence is divided into N blocks of the length n bits, then a test statistic value is computed for each block. All 15 tests may be divided into two types (according to the test statistic distribution) – the binomial-based (i. e. two-sided) and the chi-square-based (i. e. one-sided) tests. The Frequency test, the Runs test, the Discrete Fourier transform test, Maurer's universal statistical test, the

²Actually, the Non-overlapping template matching test, the Random excursions test, the Random excursions variant test, the Serial test and the Cumulative sums test generate 148, 8, 18, 2 and 2 p -values, respectively, other tests – 1 p -value each; however, it is very common considering only 1 of p -value from the set of p -values related to the test.

Random excursions variant test and the Cumulative sums test belong to the binomial-based tests, and the others are chi-square-based tests.

For each value of test statistic the p -value is computed according to the theoretic test statistic distribution³. A test is considered to be passed when the computed p -value is larger than the significance level α . According to the computed p -values for N blocks, each test performs two-level test – the first-level and the second-level tests. The first-level test focuses on the passing ratio of the N p -values, and the second-level test further focuses on the uniformity of the N p -values to assess whether the test statistic values follow the expected distribution, i. e. the chi-square distribution. The two-level testing approach was found to increase the testing capability [4].

Related works

There are many articles on the NIST SP 800-22 test suite. Among them S.-J. Kim, K. Umeno and A. Hasegawa in work [5] have found that the test setting of the Discrete Fourier transform test and Lempel – Ziv test from an earlier version of NIST SP 800-22 suite are wrong, they give four corrections of mistakes in the test settings, K. Hamano and T. Kaneko in papers [6; 7] derived the distribution function of the discrete Fourier transform spectrum, changed the threshold value from the default value of $\sqrt{3n}$ to the value of $\sqrt{\left(\ln \frac{1}{0.05}\right)n}$,

where n is the length of a random number sequence, and corrected the occurrence probabilities π_i of the Overlapping template matching test included. F. Sulak, A. Doğanaksoy, B. Ege and O. Koçak in article [8] found that the p -values for short sequences (less than 512 bits) follow a specific discrete distribution, rather than the assumed uniform distribution for long sequences, calculated the probabilities of subintervals for some NIST test and proposed an alternative approach to evaluate test results for short sequences. F. Pareschi, R. Rovatti and G. Setti in work [4] investigated the reliability of the second-level tests, and analysed the sensitivity to the approximation errors introduced by the computation of p -values. Since the sequence length is finite in practice and thus the set of possible statistic values is discrete, F. Pareschi et al. in paper [9] provided the more closer distributions to the actual ones of p -values for the Frequency test, the Runs test, the Discrete Fourier transform test.

Recently, Yu. S. Kharin and A. M. Zubkov in article [10] proposed a model of a complex null hypothesis allowing possible deviations for the distribution of output sequences of generators from the equiprobable distribution, and developed an approach to the construction of the criterion for testing such hypothesis. The results obtained in work [10] may be applied in practice in statistical testing of physical random sequence generators.

First-level testing approach of the NIST SP 800-22

Denote by $\mathcal{H}_0$ the null hypothesis on the uniformity and independence of binary sequence elements, i. e. that elements of binary sequence are independent and take values 0 and 1 with probabilities $\frac{1}{2}$. In other words, hypothesis $\mathcal{H}_0$ is that the RNG under test produces sequences with the above properties, such a generator will be called an ideal RNG further. In the statistical hypothesis testing, a test statistic is chosen and used to determine whether $\mathcal{H}_0$ should be accepted or rejected. Under the null hypothesis $\mathcal{H}_0$, the theoretical reference distribution of the statistic is determined by the mathematical methods. From the reference distribution, a range of acceptable statistic values is determined based on a preset coefficient $\gamma = 1 - \alpha$ (for example, $\gamma = 0.99$), where α is the significance level, i. e. γ is the probability that the statistic value is inside of the acceptable statistic value range. If the test statistic value lies outside the acceptable statistic value range, the hypothesis $\mathcal{H}_0$ is rejected, otherwise, the hypothesis $\mathcal{H}_0$ is accepted.

A randomness test suite may contain a serial of tests, which evaluate different properties of sequences. Since the ranges of acceptable values of various test statistics corresponding to the same γ are different, the p -values p , determined from the distributions of test statistics, may be used as unified numerical characteristics of the test results for sequences. In publication [2] the p -value is defined as «the probability that a perfect RNG would have produced a sequence less random than the sequence that was tested» [2, p. 1-4]. More specifically, the p -value is the probability of obtaining a statistic value s equal to or «more extreme» than the observed statistic value s_{obs} for the tested sequence, under the assumption that the null hypothesis is correct. According to the definition of «more extreme» cases, the tests are generally divided into two categories – one-sided tests and two-sided tests.

When $\mathcal{H}_0$ is true and $p \leq \alpha$ we have a false positive in the test interpretation. This is called type I error, we can compute its probability since we have a complete characterisation of the sequences generated by a TRNG.

³Remind that the p -value is the probability to obtain the value of statistic at least as large as the actually observed one, under the assumption that the null hypothesis is true.

It is known that if the distribution of statistic is continuous then p -values should be uniformly distributed on the interval $[0, 1]$ under the null hypothesis. Since p is uniformly distributed, the probability of a type I error is

$$\mathbf{P}_{\mathcal{H}_0}\{p \leq \alpha\} = F_p^{(0)}(\alpha) = \alpha,$$

where $F_p^{(0)}$ is the distribution function of the random variable p having a uniform distribution on the interval $[0, 1]$. For this reason, α is also called level of significance or statistical significance. The value of α is usually small, the significance level recommended by NIST is $\alpha = 0.01$.

On the contrary, the fact that $p > \alpha$ when $\mathcal{H}_0$ is false we have a false negative in the test interpretation. This is called type II error (accepting the sequence as random when $\mathcal{H}_0$ is false), and we denote this probability by β . The value $(1 - \beta)$ is called the statistical power of the test, its exact computation is a difficult task, since it depends on the specific probabilistic mathematical model of the (non-ideal) generator under test.

So we can commit two errors:

- type I – reject $\mathcal{H}_0$, when the sequence is generated by an ideal RNG;
- type II – accept $\mathcal{H}_0$, when the sequence is generated by a RNG that is non-ideal.

It's important to find a balance between the probabilities of making type I and type II errors. Reducing α always comes at the cost of increasing β , and vice versa, with a fixed sample size. This approach is known as first-level testing and is followed by paper [4].

Second-level testing approach of the NIST SP 800-22

The usual way to test a TRNG or PRNG is to generate a sequence of n bits and analyse it with the test suite. Given a level of significance α , the sequence may be considered random if all tests in the suite produce p -values greater than α .

This approach presents a serious weakness. It is well known that some PRNGs can easily pass all tests. For example, a periodic (and thus, not ideal) generator always passes the Frequency test if the number of ones and zeros in the sequence is balanced.

To overcome this weakness, a more intensive test is necessary, involving a number N of different sequences generated by the RNG under test. In publication [2] NIST recommends using the second-level testing approach: a long binary sequence is partitioned into N subsequences, each with n bits. A standard test is applied for each sequence, and the distribution of the N obtained p -values is compared with a uniform distribution $F_p^{(0)}$. Note that from the definition of null hypothesis it follows that if $\mathcal{H}_0$ is true, then the N sequences (and so the N p -values) are independent of one another and approximately uniformly distributed on the interval $[0, 1]$. To check this in publication [2] NIST proposes a chi-square goodness-of-fit test, this test is again a statistical test and gives another (a second-level) p -value p_{II} .

The NIST SP 800-22 suite includes 15 tests and for each of them the second-level test is performed. For each test and for each of N blocks of the sequence N p -values are computed, then the following second-level test is performed based on these N p -values.

This approach has been known for a long time [3] and it may increase the testing power as compared to a standard approach [4; 9].

1. Given the empirical results for a particular statistical test, compute the fraction ζ of the sequences passed the test. For example, if N binary sequences were tested with the significance level $\alpha = 0.01$ and m binary sequences had p -values equal to or greater than α , then the fraction is $\zeta = \frac{m}{N}$. If $\mathcal{H}_0$ is true, then the distribution of ζN is binomial with parameters N and $(1 - \alpha)$; if N is large enough (e. g., $N \geq 1000$), then the distribution of ζ may be approximated by the distribution of the random variable having a normal distribution $\mathcal{N}(\mu, \sigma)$, where

$$\mu = 1 - \alpha; \sigma = \sqrt{\frac{\alpha(1 - \alpha)}{N}}.$$

The range of acceptable fractions may be determined by the interval

$$\left[1 - \alpha - 3\sqrt{\frac{\alpha(1 - \alpha)}{N}}, 1 - \alpha + 3\sqrt{\frac{\alpha(1 - \alpha)}{N}} \right].$$

If the ζ falls outside of this interval, then there is evidence that the data does not conform to the null hypothesis and $\mathcal{H}_0$ is rejected. The normal distribution approximation for the binomial distribution is reasonably accurate for large sample sizes.

2. Given N sequences, the distribution of p -values is examined to check the uniformity. The interval $[0, 1]$ is divided into k disjoint subintervals, and the number of p -values falling in each subinterval are counted. Uniformity is determined via an application of a chi-square goodness-of-fit test in k subinterval with statistic χ^2 . This is accomplished by computing chi-square statistic value

$$\chi_{\text{obs}}^2 = \sum_{i=1}^k \frac{\left(\pi_i - \frac{N}{k}\right)^2}{\frac{N}{k}},$$

where π_i is the number of p -values in i^{th} subinterval. This statistical test yields a level-two p -value $p_{\Pi} = \mathbf{P}_{\mathcal{H}_0}\{\chi^2 > \chi_{\text{obs}}^2\}$ (here χ^2 is a random variable having chi-square distribution with $(k-1)$ degrees of freedom), which is calculated as follows:

$$p_{\Pi} = \frac{1}{\Gamma\left(\frac{k-1}{2}\right)} \int_{\frac{\chi_{\text{obs}}^2}{2}}^{\infty} t^{\frac{k-1}{2}-1} e^{-t} dt, \quad \Gamma\left(\frac{k-1}{2}\right) = \int_0^{\infty} t^{\frac{k-1}{2}-1} e^{-t} dt.$$

Given a significance α_{Π} , $\mathcal{H}_0$ is rejected if $p_{\Pi} \leq \alpha_{\Pi}$, otherwise the sequences may be considered to be uniformly distributed. In publication [2] NIST recommends to use $\alpha_{\Pi} = 0.0001$ and $k = 10$ subintervals.

The choice of N in a two-level approach is usually a trade-off. For example, the length of the sequence n may be limited by the memory size or by the computational power available. On the other hand, it may be preferable to test short sequences, thereby limiting n .

In every statistical test some approximations are adopted, introducing errors in the p -value computation and so in the p -value distribution. It was observed [4; 9] that for extremely large values of N , the level-two approach always fail, it ends with $p_{\Pi} \approx 0$. In this case we can say that the test is not reliable, since its significance is sensible different from the expected one. From this point of views, N is also a trade-off between statistical power $(1-\beta)$ and reliability of a level-two test.

Advantages of two-level test

The undoubted advantage of using the two-level test, for example, is rejection the generator with a small period size by Frequency test: regardless of the sequence length, a basic first-level test is always passed, while the second-level test is able to recognise a generator that always passes a basic test is not ideal.

Two PRNGs were considered in work [4]: the 32-bit version of the KISS [11], which is a very simple but effective generator, and the BBS generator [12]:

$$x_{n+1} = x_n^2 \bmod M, \quad M = P \cdot Q,$$

where P and Q are large distinct primes; x_0 is a quadratic residue mod M ; the output of the BBS generator is the bit parity of x_{n+1} , that is a computationally complex PRNG that, under the assumption of intractability for quadratic residuosity problem modulo M , has proven to be cryptographically secure (i. e. polynomial-time unpredictable (see [12])). For both generators in work [4] were performed a first-level test on a single sequence and a second-level test, checking the uniformity of $N = 10\,000$ p -values, on N different sequences with a chi-square test over 16 subintervals, and with a Kolmogorov – Smirnov test.

The empirical distributions of the tests of p -values were computed for both generators, which were then compared with the uniform distribution in the interval $[0, 1]$ by calculating the p -values p_{Π} . In this case, $\mathcal{H}_0$ correspond to «the two distributions match»; again, we reject $\mathcal{H}_0$ if $p_{\Pi} \leq \alpha_{\Pi}$ and accept $\mathcal{H}_0$ if $p_{\Pi} > \alpha_{\Pi}$. Although in publication [2] NIST recommends to use $\alpha_{\Pi} = 0.0001$, the value of α_{Π} in work [4] was set to $\alpha = 0.01$ to make possible a direct comparison between a first-level test and a second-level test. The authors noted that the comparison may seem unfair, since a first level-test considers n bits, while a second-level test – nN bits.

Both generators had passed all first-level tests. However (except for the Discrete Fourier transform test that is well-known to have an error in the parameter of the reference distribution of the considered statistic), only the BBS generator passed the second-level tests. The proposed second-level test is able to recognise the non-randomness of the KISS generator, while a simple first-level test fails. In addition to the above experiment, the authors of work [4] considered a second one involving a much larger number of sequences generated with the BBS algorithm ($N = 150\,000$ sequences) in order to obtain more reliable results. The tests were also

applied to three high-end physical process based TRNGs. All three generators have been considered with an additional post processing stage, this in order to hide all possible imperfections and be sure that the properties of analysing streams are close to the properties of sequences of independent and equiprobable bits. Results of experiments are far from the desired ones, since too many tests fail, in particular, only 6 out of 15 two-level tests were passed by the BBS generator.

Experiments

All 15 tests from table 1 were applied to pseudorandom sequences generated by AES block cipher. The detailed description of the design of AES-based RNG may be found in papers [13; 14]. For generation such sequences AES block cipher in the CFB (*cipher feedback block*) mode (fig. 1) with the zero initialisation vector (IV) and plaintext block (Plaintext) was used. The key for each sequence was randomly selected from the set of 128-bit binary sequences, the length in bits of all sequences was chosen to be the same and equal to $n = 2^{20} = 1\,048\,576$ each (value corresponds to the NIST input size recommendation for the length of tested sequences).

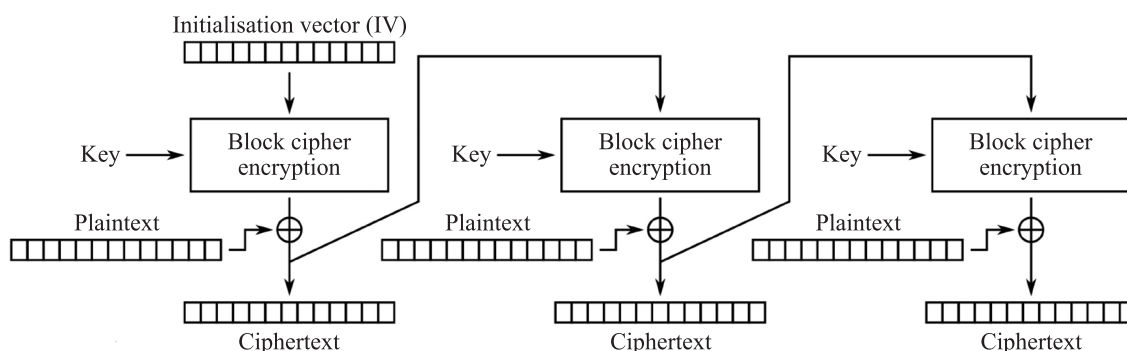


Fig. 1. CFB mode

Keys (on fig. 1) were obtained using the open source cryptographic library OpenSSL:

```
int RAND_bytes(unsigned char *buf, int num).
```

This function generates *num* bytes using the cryptographically secure PRNG (CSPRNG), and saves them in the *buf* array. By default, the OpenSSL CSPRNG supports a security level of 256 bits, provided it was able to seed itself from a trusted entropy source⁴.

If we will use the AES-based PRNG, then this allow us to arbitrarily increase N . The effect of increasing the testing power can be seen in the example of table 2, where the testing results for the AES-based RNG are shown. The two-level test was performed for all 188 statistics values computed by NIST test suite, but only 15 of them are presented in table 2. In order to have the same significance in all tests, α_{II} in all experiments was set equal to $\alpha = 0.01$. These results confirm that for extreme values of N (e. g., $N = 2^{20}$) the two-level testing approach was carried out on the sequences obtained by AES algorithm (it is known that such sequences are practically indistinguishable from random one (see [13; 14])) are failed too, i. e. it ends with $p_{II} \approx 0$.

Table 2

Results of the chi-square-based two-level randomness test
for the AES-based RNG, with N ranging from 10^3 to 2^{20}

Number	Test name	Results of the chi-square test			
		$N = 10^3$	$N = 10^4$	$N = 10^5$	$N = 2^{20}$
1	Frequency test	0.616 305	0.290 806	0.588 411	0.000 803
2	Block frequency test	0.187 581	0.773 212	0.374 097	0.000 125
3	Cumulative sums test	0.401 199	0.124 765	0.959 543	0.000 009
4	Runs test	0.150 340	0.885 418	0.910 568	0.107 966
5	Longest run test	0.610 070	0.239 883	0.000 355	0.000 000

⁴RAND_bytes [Electronic resource]. URL: https://www.openssl.org/docs/man1.1.1/man3/RAND_bytes.html (date of access: 21.06.2025).

Ending of the table 2

Number	Test name	Results of the chi-square test			
		$N = 10^3$	$N = 10^4$	$N = 10^5$	$N = 2^{20}$
6	Binary matrix rank test	0.878 618	0.341 017	0.000 000	0.000 000
7	Discrete Fourier transform test	0.371 941	0.014 836	0.000 000	0.000 000
8	Overlapping template matching test	0.071 177	0.202 268	0.000 000	0.000 000
9	Maurer's universal statistical test	0.574 903	0.108 534	0.000 000	0.000 000
10	Approximate entropy test	0.246 750	0.078 038	0.219 501	0.000 000
11	Serial test	0.942 198	0.174 057	0.213 964	0.572 679
12	Linear complexity test	0.839 507	0.279 152	0.299 852	0.117 305
13	Non-overlapping template matching test	0.092 041	0.372 782	0.121 382	0.275 416
14	Random excursions test	0.914 727	0.663 838	0.346 173	0.000 028
15	Random excursions variant test	0.238 264	0.133 576	0.000 080	0.000 000

Note. Tests, where $p_{\Pi} \leq 0.01$, are in bold.

In order to identify the problem, let's take a look on the simple Frequency test.

Frequency test

The purpose of the Frequency test is to determine whether the number of ones and zeros in a sequence are approximately the same as would be expected for a truly random sequence. The test assesses the closeness of the fraction of ones to $\frac{1}{2}$, that is, the number of ones and zeros in a sequence should be about the same. All subsequent tests depend on passing this first basic test (see [2, p. 2-2]).

Let n be the length of the bit string, $\varepsilon = \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ be the input sequence of bits and the null hypothesis $\mathcal{H}_0$ is that the sequence ε consists of independent identically distributed Bernoulli random variables $(X_i \text{ or } \varepsilon_i)$, where $X_i = 2\varepsilon_i - 1$, and so the probability of ones is $\frac{1}{2}$. Denote by

$$S_n = X_1 + \dots + X_n = 2(\varepsilon_1 + \dots + \varepsilon_n) - n.$$

Under the null hypothesis, S_n is assumed to follow a binomial distribution. By the classic de Moivre – Laplace theorem, for a sufficiently large number of trials, the distribution of the binomial sum, normalised by $\sqrt{n}$, is closely approximated by a standard normal distribution. This test makes use of that approximation to assess the closeness of the fraction of ones to $\frac{1}{2}$.

The test is derived from the well-known central limit theorem for the random walk, $S_n = X_1 + \dots + X_n$. According to the central limit theorem,

$$\lim_{n \rightarrow \infty} F_{S_n}(z) = \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{u^2}{2}} du,$$

where $F_{S_n}(z) = \mathbf{P}\left\{\frac{S_n}{\sqrt{n}} < z\right\}$. This classical result serves as the basis of the simplest test for randomness. It implies that, if S_n is normal, then $|S_n|$ is half normal (i. e. $F_{|S_n|}(z) = 2F_{S_n}(z)$, $z \geq 0$) and

$$F_{|S_n|}(z) = \mathbf{P}\left\{\frac{|S_n|}{\sqrt{n}} \leq z\right\}, \quad \lim_{n \rightarrow \infty} F_{|S_n|}(z) = 2\Phi(z) - 1.$$

So, according to the test statistic $S = \frac{|S_n|}{\sqrt{n}}$, it is necessary to determine observed value $S(\varepsilon) = \frac{|X_1 + \dots + X_n|}{\sqrt{n}}$, and then calculate the corresponding p -value, which is equal to

$$p_1 = \lim_{n \rightarrow \infty} \left(1 - F_{|S_n|}(S(\varepsilon)) \right) = 2 \left(1 - \Phi(S(\varepsilon)) \right) = \operatorname{erfc} \left(\frac{S(\varepsilon)}{\sqrt{2}} \right),$$

where $\operatorname{erfc}(\cdot)$ is the complementary error function

$$\operatorname{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^{\infty} e^{-u^2} du.$$

In fig. 2, *a*, is shown the histogram of the set of p -values for $k = 10$ subintervals corresponding to the Frequency test applied to the AES-based generator, and for the comparison in fig. 2, *b*, – the Binary matrix rank test from the NIST test suite applied to the same sequences from the AES-based generator. If we consider the theoretical standard deviation of the number of random $N = 2^{20}$ independent, uniformly distributed values over $k = 10$ subintervals, we easily get $\sigma = \frac{\sqrt{N(k-1)}}{k}$. In this case, $\sigma \approx 307.2$, the upper and lower continuous lines in the figures correspond to the deviation of 3σ from the mean value; the figure shows that the observed deviation is far from this value.

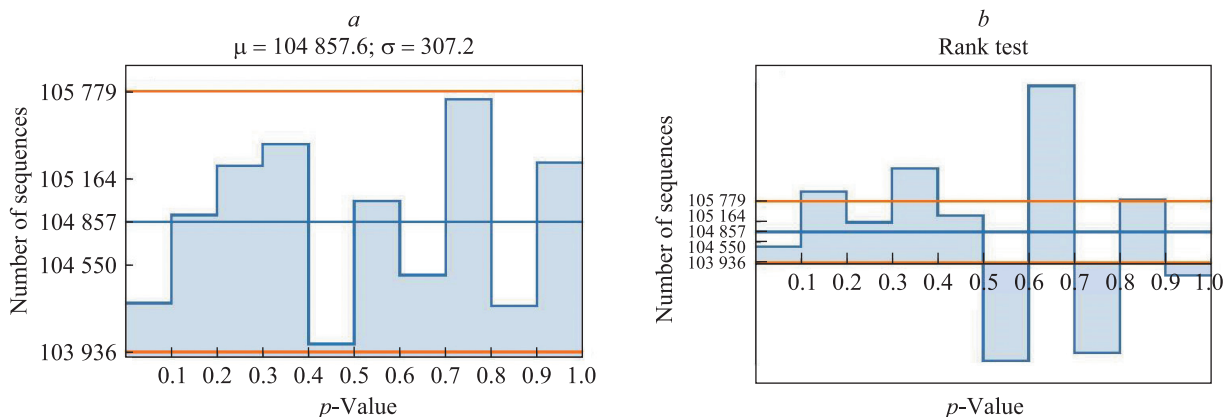


Fig. 2. Comparison between expected deviation and measured deviation in the distribution of p -values in the interval $[0, 1]$ for the Frequency test (*a*) and the Binary matrix rank test (*b*)

Let $p_1, \dots, p_N \in [0, 1]$ be the sample of p -values with assumed continuous CDF $F_p^{(0)}(x)$. The empirical CDF $F_N(x)$ for this sample is defined as follows:

$$F_N(x) = \frac{1}{N} \sum_{j=1}^N H(x - p_j),$$

where $H(x)$ is the unity-step function, defined as $H(x) = \begin{cases} 1, & x \geq 0, \\ 0, & x < 0. \end{cases}$ The Kolmogorov – Smirnov statistic for a given CDF $F_p^{(0)}(x)$ is

$$D_N = \sup_x |F_N(x) - F_p^{(0)}(x)|.$$

Let $\mathcal{K}$ be the CDF of D_N , then the p -values is given by $1 - \mathcal{K}(D_N)$. Under null hypothesis that the sample comes from the hypothesised distribution $F(x)$

$$\sqrt{N}D_N \xrightarrow[N \rightarrow \infty]{} \sup_t |B(F_p^{(0)}(t))|$$

in distribution, where $B(t)$ is the Brownian bridge.

In fig. 3, *a*, is shown the discrete empirical CDF F_N of p -values for the Frequency test applied to the AES-based sequences, in fig. 3, *b*, are presented the differences between empirical CDF F_N and uniform CDF, $n = N = 2^{20}$, from where it is clear that the modulus of the difference between F_N and $F_p^{(0)}$ does not exceed 0.001 2.

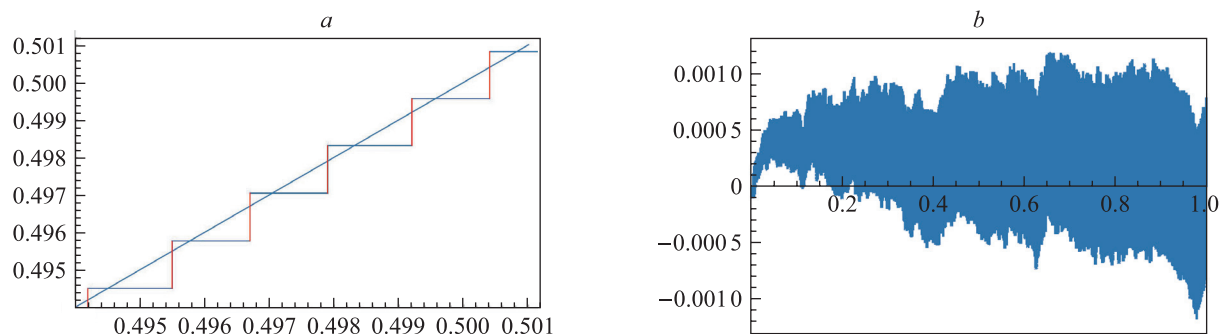


Fig. 3. Comparison between empirical CDF F_N of the p -values generated by Frequency test for $n = 2^{20}$ and uniform CDF (a). Deviation of the empirical CDF F_N of the p -value distribution of the Frequency test from continuous uniform CDF (b)

The deviation may be identified with an error propagated from the computation of the p -value in the first-level test, which is a consequence of the introduced approximations. The random variable S_n in relation (1) has a binomial distribution, which is however assumed to be normal. From Berry – Esseen theorem, we know that the error of this approximation, under the assumptions of the central limit theorem, is bounded by (see [15])

$$\sup_x |F_{S_n}(x) - \Phi(x)| \leq \frac{CE|X_i|^3}{\sigma^3\sqrt{n}},$$

where $F_{S_n}(z) = \mathbf{P}\left\{\frac{S_n}{\sqrt{n}} < z\right\}$; $\Phi(\cdot)$ is a CDF of the standard normal distribution; n is the number of independent variables X_i , $1 \leq i \leq n$, summed in relation (1), i. e. the number of bits in the sequence; $\sigma = \mathbf{E}(X_i^2) = 1$; the third order moment is $\mathbf{E}|X_i|^3 = 1$; $C < 0.4748$ (see [16]). The maximal error ϵ of p -value is the error of the approximation $F_{S_n}(x)$ by $\Phi(x)$ and from formula (2) is twice the above error, i. e. $\epsilon = \frac{2C}{\sqrt{n}}$. If $n = 2^{20}$, then $\epsilon = 9.3 \cdot 10^{-4}$.

Assuming this bound of the error in the computation of a p -value, we can bound also the maximal error in the distribution of N p -values in k subintervals. A p -value that should belong to a subinterval may be found into the nearby one only if its distance from the border of two subintervals is less than ϵ . If we have N p -values uniformly distributed in the interval $[0, 1]$ the maximal number of p -values that may be found in the wrong adjacent subinterval is ϵN . This is independent of the numbers of subintervals.

Since all subintervals (but the first and the last), have two neighbours, the maximum error Δ in the number of p -values in a subinterval is $\Delta = 2N\epsilon$. In our case, $N = 2^{20}$, so $\Delta \approx 1944.8$. This value is compatible with that may be observe in fig. 2, a. The increasing n will result in decreasing the propagated error as $\frac{1}{\sqrt{n}}$ (the results of the experiments are confirm that in paper [4]). More generally, increasing n will reduce the error in the first-level p -value and so increasing the reliability of a second-level test, for all tests in the suite. A very simple reliability condition is requiring that Δ is smaller than the standard deviation of the number of p -values in a subinterval from its mean value, i. e.

$$\Delta < \frac{\sqrt{N(k-1)}}{k}.$$

In the case $\Delta = 2N\epsilon$, $k = 10$ and $n = 2^{20}$, we get

$$N \leq \frac{1}{4\epsilon^2 k} \left(1 - \frac{1}{k}\right) \approx 26\,014.6.$$

Two-sided estimates for the quantiles of the binomial distribution

To construct a statistical test with the specified error probabilities, we need the information on the distribution «tails». But relative errors of the de Moivre – Laplace approximations for the tails of binomial distribution

function are large. This was one of the first problems in the theory of large deviations. There are a lot of inequalities for the tails of binomial distribution function, as well as for the distributions of sums of random variables.

In paper [17] explicit two-sided estimates for the distribution function $F_{n,p}(k)$ of the binomial law with parameters n, p were obtained. These results are close in form to the Bernstein and Feller inequalities, but have an explicit form and complement them, almost completely solving the problem of estimating the probabilities of moderate and large deviations for binomial laws for any values of the parameters.

Let $X_{n,p}$ be a random variable with the binomial distribution with parameters n, p :

$$\mathbf{P}\{X_{n,p} \leq k\} = \sum_{0 \leq m \leq k} C_n^m p^m (1-p)^{n-m}.$$

Theorem 1 [17]. *Let*

$$H(x, p) = x \ln \left(\frac{x}{p} \right) + (1-x) \ln \left(\frac{1-x}{1-p} \right), \quad \text{sign}(x) = \frac{x}{|x|}$$

for $x \neq 0$ and $\text{sign}(0) = 0$, and let $\{C_{n,p}(k)\}_{k=0}^n$ be increasing sequences defined as follows:

$$C_{n,p}(0) = (1-p)^n, \quad C_{n,p}(n) = 1-p^n,$$

$$C_{n,p}(k) = \Phi \left(\text{sign} \left(\frac{k}{n} - p \right) \sqrt{2nH \left(\frac{k}{n}, p \right)} \right), \quad 1 \leq k < n.$$

Then for every $k = 0, 1, \dots, n-1$ and for every $p \in (0, 1)$

$$C_{n,p}(k) \leq \mathbf{P}\{X_{n,p} \leq k\} \leq C_{n,p}(k+1)$$

and equalities hold only for $k = 0$ and $k = n-1$.

We will use the following standard notations:

$$\lfloor x \rfloor = \max \{m \in \mathbb{Z} \mid m \leq x\}, \quad \lceil x \rceil = \min \{n \in \mathbb{Z} \mid n \geq x\}.$$

By definition of the α -level quantile $x_\alpha(n, p)$ of the binomial distribution with parameters n, p

$$x_\alpha(n, p) = \min \{k : \mathbf{P}\{X_{n,p} \leq k\} \geq \alpha\} \in \{0, 1, \dots, n\},$$

$$x_0(n, p) = 0 \quad \text{and} \quad x_1(n, p) = n,$$

$$x_{\frac{1}{2}} \left(n, \frac{1}{2} \right) = \left\lfloor \frac{n}{2} \right\rfloor.$$

From theorem 1 for any $\alpha \in [0, 1]$ it follows that

$$\min \{k : C_{n,p}(k+1) \geq \alpha\} \leq x_\alpha(n, p) = \min \{k : \mathbf{P}\{X_{n,p} \leq k\} \geq \alpha\} \leq \min \{k : C_{n,p}(k) \geq \alpha\}.$$

Recently the author had obtained the two-sided estimates for the α -level quantile $x_\alpha(n, p)$ of the binomial law. Such estimates allow for any α -level and any values of binomial law parameters n, p to determine a narrow integer interval containing the quantile $x_\alpha(n, p)$, it has been experimentally shown that quite often the length of this interval is one, i. e. the interval boundaries are two subsequent integers one of which is the quantile $x_\alpha(n, p)$. These estimates allow to completely eliminate the systematic error in determining one of the $k = 10$ disjoint subintervals on the interval $[0, 1]$ for the Frequency test to which the p -value belongs.

Theorem 2 [18]. *The following estimates are true:*

- for $0 < \alpha < \frac{1}{2}$

$$r_1 - 1 \leq x_\alpha(n, p) \leq r_2,$$

where $q = 1 - p$, $r_1 = \left\lceil np + \frac{q-p}{6} \Phi^{-1}(\alpha)^2 + \Phi^{-1}(\alpha) \sqrt{npq + \frac{(q-p)^2 \Phi^{-1}(\alpha)^2}{6^2}} \right\rceil$, and

$$r_2 = \left\lceil \frac{(6-5p)np + \frac{q}{2} \Phi^{-1}(1-\alpha) \left(3q \Phi^{-1}(1-\alpha) - \sqrt{24(6-5p)np + (3+p)^2 \Phi^{-1}(1-\alpha)^2} \right)}{(6-5p) + \frac{2q \Phi^{-1}(1-\alpha)^2}{n}} \right\rceil;$$

• for $\frac{1}{2} < \alpha \leq 1$

$$r_3 \leq x_\alpha(n, p) \leq r_1,$$

where $r_3 = \left\lceil \frac{(1+5p)np + \frac{p}{2} \Phi^{-1}(\alpha) \left((4-3p) \Phi^{-1}(\alpha) + \sqrt{24n(1+5p)q + \Phi^{-1}(\alpha)^2 (4-p)^2} \right)}{(1+5p) + \frac{2p \Phi^{-1}(\alpha)^2}{n}} \right\rceil - 1;$

• for $\alpha = \frac{1}{2}$

$$x_{\frac{1}{2}}(n, p) = \begin{cases} \lfloor np \rfloor \text{ or } \lceil np \rceil, & \text{if } p \neq \frac{1}{2}, \\ \frac{n-1}{2}, & \text{if } p = \frac{1}{2}, n - \text{odd}, \\ \frac{n}{2}, & \text{if } p = \frac{1}{2}, n - \text{even}. \end{cases}$$

Corollary. If $0 < \alpha < \frac{1}{2}$, then the following estimates are true:

$$\left\lfloor \frac{n}{2} + \frac{\sqrt{n}}{2} \Phi^{-1}(\alpha) \right\rfloor - 1 \leq x_\alpha\left(n, \frac{1}{2}\right) \leq \left\lceil \frac{\frac{7n}{4} + \frac{3}{8} \Phi^{-1}(\alpha)^2 + \frac{1}{4} \Phi^{-1}(\alpha) \sqrt{42n + \left(\frac{7}{2}\right)^2 \Phi^{-1}(\alpha)^2}}{\frac{7}{2} + \frac{\Phi^{-1}(\alpha)^2}{n}} \right\rceil,$$

if $\frac{1}{2} < \alpha < 1$, then

$$\left\lceil \frac{\frac{7n}{4} + \frac{5}{8} \Phi^{-1}(\alpha)^2 + \frac{1}{4} \Phi^{-1}(\alpha) \sqrt{42n + \left(\frac{7}{2}\right)^2 \Phi^{-1}(\alpha)^2}}{\frac{7}{2} + \frac{\Phi^{-1}(\alpha)^2}{n}} \right\rceil - 1 \leq x_\alpha\left(n, \frac{1}{2}\right) \leq \left\lfloor \frac{n}{2} + \frac{\sqrt{n}}{2} \Phi^{-1}(\alpha) \right\rfloor,$$

if $\alpha = \frac{1}{2}$, then

$$x_{\frac{1}{2}}\left(n, \frac{1}{2}\right) = \begin{cases} \frac{n-1}{2}, & \text{if } n - \text{odd}, \\ \frac{n}{2}, & \text{if } n - \text{even}. \end{cases}$$

The 11 quantiles of binomial distribution with parameters $\left(2^{20}, \frac{1}{2}\right)$ are shown in table 3.

Table 3

Quantiles of binomial distribution

α	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
$x_{\alpha}\left(2^{20}, \frac{1}{2}\right)$	0	523 632	523 857	524 020	524 158	524 288	524 418	524 556	524 719	524 944	2^{20}

Conclusions

The two-level approach for testing RNGs involving the well-known NIST SP 800-22 test suite was considered. Such approach may increase the reliability of the test. However, it is sensitive to the approximation error introduced by the computing of p -values. In this paper it has been shown that for AES-based sequences two-level testing approach is not reliable too. Systematic error in the computing of the p -values is dependent only on n , that is the number of bits in the analysed sequences. For a reliable second-level test, this error, based on the carried out heuristic considerations, should be no greater than $\frac{\sigma}{N} = \frac{1}{k} \sqrt{\frac{k-1}{N}}$, where σ is the standard deviation from the mean number of particles in a bin for equiprobable scheme of allocation N particles in k bins (disjoint subintervals of the interval $[0, 1]$). For example, in the second-level test of the Frequency test with $n = 2^{20}$ the number of tested sequences N should not exceed 26 000.

To completely eliminate the systematic error appearing in the Frequency test when determining the number of bin from k bins to which the p -value belongs the two-sided estimates of the quantiles of the binomial law are proposed.

References

1. Menezes AJ, van Oorschot PC, Vanstone SA. *Handbook of applied cryptography*. Boca Raton: CRC Press; 1996. XXVIII, 780 p. (Discrete mathematics and its applications).
2. Rukhin A, Soto J, Nechvatal J, Smid M, Barker E, Leigh S, et al. *A statistical test suite for random and pseudorandom number generators for cryptographic applications*. Gaithersburg: National Institute of Standards and Technology; 2010. [131] p. (NIST special publications; SP 800-22 revision 1a).
3. L'Ecuier P, Simard R. *TestU01: a software library in ANSI C for empirical testing of random number generators* [Internet]. Montreal: Université de Montréal; 2013 [cited 2025 June 21]. IV, 214 p. Available from: <https://scispace.com/pdf/a-software-library-in-ansi-c-for-empirical-testing-of-random-hk10yxs2w3.pdf>.
4. Pareschi F, Rovatti R, Setti G. Second-level NIST randomness tests for improving test reliability. In: Institute of Electrical and Electronics Engineers. *2007 IEEE International symposium on circuits and systems (ISCAS); 2007 May 27–30; New Orleans, USA*. [S. l.]: Institute of Electrical and Electronics Engineers; 2007. p. 1437–1440. DOI: 10.1109/ISCAS.2007.378572.
5. Kim S-J, Umeno K, Hasegawa A. *Corrections of the NIST statistical test suite for randomness*. arXiv:nlin/0401040v1 [Pre-print]. 2004 [cited 2025 June 21]: [14 p.]. Available from: <https://arxiv.org/abs/nlin/0401040v1>.
6. Hamano K. The distribution of the spectrum for the Discrete Fourier transform test included in SP 800-22. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*. 2005;E88-A(1):67–73.
7. Hamano K, Kaneko T. Correction of overlapping template matching test included in NIST randomness test suite. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*. 2007;E90-A(9):1788–1792.
8. Sulak F, Doğanaksoy A, Ege B, Koçak O. Evaluation of randomness test results for short sequences. In: Carlet C, Pott A, editors. *Sequences and their applications – SETA 2010. Proceedings of the 6th International conference; 2010 September 13–17; Paris, France*. Berlin: Springer; 2010. p. 309–319 (Lecture notes in computer science; volume 6338). DOI: 10.1007/978-3-642-15874-2_27.
9. Pareschi F, Rovatti R, Setti G. On statistical tests for randomness included in the NIST SP 800-22 test suite and based on the binomial distribution. *IEEE Transactions on Information Forensics and Security*. 2012;7(2):491–505. DOI: 10.1109/TIFS.2012.2185227.
10. Kharin YuS, Zubkov AM. On statistical testing of composite hypotheses on s -dimensional uniform probability distribution of binary sequences. *Diskretnaya matematika*. 2024;36(1):116–135. Russian. DOI: 10.4213/dm1806.
11. Rose GG. KISS: a bit too simple. *Cryptography and Communications*. 2018;10(1):123–137. DOI: 10.1007/s12095-017-0225-x.
12. Blum L, Blum M, Shub M. A simple unpredictable pseudo-random number generator. *SIAM Journal on Computing*. 1986; 15(2):364–383. DOI: 10.1137/0215025.
13. Zubkov AM, Serov AA. Testing the NIST statistical test suite on artificial pseudorandom sequences. *Matematicheskie voprosy kriptografii*. 2019;10(2):89–96. DOI: 10.4213/mvk286.

14. Zubkov AM, Serov AA. A natural approach to the experimental study of dependence between statistical tests. *Matematicheskie voprosy kriptografii*. 2021;12(1):131–142. DOI: 10.4213/mvk352.
15. Shiryaev AN. *Probability*. 2nd edition. Boas RP, translator. Berlin: Springer-Verlag; 1996. XVI, 623 p. (Graduate texts in mathematics; volume 95).
16. Korolev V, Shevtsova I. An improvement of the Berry – Esseen inequality with applications to Poisson and mixed Poisson random sums. *Scandinavian Actuarial Journal*. 2012;2012(2):81–105. DOI: 10.1080/03461238.2010.485370.
17. Zubkov AM, Serov AA. A complete proof of universal inequalities for distribution function of binomial law. *Theory of Probability and its Applications*. 2013;57(3):539–544. DOI: 10.1137/S0040585X97986138.
18. Serov AA. Two-sided estimates of the binomial law quantiles. *Theory of Probability and its Applications*. Forthcoming 2026.

Received 24.12.2025 / revised 19.01.2026 / accepted 29.01.2026.

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

PROBABILITY THEORY AND MATHEMATICAL STATISTICS

УДК 519.214

ПРЕДЕЛЬНАЯ ТЕОРЕМА О СХОДИМОСТИ К ЛОКАЛЬНОМУ ВРЕМЕНИ БРОУНОВСКОЙ ЭКСКУРСИИ

В. И. АФАНАСЬЕВ¹⁾

¹⁾Математический институт им. В. А. Стеклова РАН,
ул. Губкина, 8, 119333, г. Москва, Россия

Аннотация. Исследуется целочисленное случайное блуждание $\{S_i, i \geq 0\}$ с нулевым сносом и конечной дисперсией σ^2 . Для случайного процесса, сопоставляющего переменной $u > 0$ число попаданий указанного блуждания в состояние $\lfloor u\sigma\sqrt{n} \rfloor$ до момента n и рассматриваемого при условии, что $S_1 > 0, \dots, S_{n-1} > 0, S_n \leq 0$, доказывается функциональная предельная теорема о его сходимости к локальному времени броуновской экскурсии.

Ключевые слова: случайное блуждание; условное броуновское движение; локальное время условных броуновских движений; функциональная предельная теорема.

Образец цитирования:

Афанасьев В.И. Предельная теорема о сходимости к локальному времени броуновской экскурсии. *Журнал Белорусского государственного университета. Математика. Информатика*. 2026;1:36–52 (на англ.).
EDN: FXNQHJ

For citation:

Afanasyev V.I. Limit theorem on convergence to the local time of a Brownian excursion. *Journal of the Belarusian State University. Mathematics and Informatics*. 2026;1:36–52.
EDN: FXNQHJ

Автор:

Валерий Иванович Афанасьев – доктор физико-математических наук; ведущий научный сотрудник отдела дискретной математики.

Author:

Valeriy I. Afanasyev, doctor of science (physics and mathematics); leading researcher at the department of discrete mathematics.
viafan@mail.ru

LIMIT THEOREM ON CONVERGENCE TO THE LOCAL TIME OF A BROWNIAN EXCURSION

V. I. AFANASYEV^a

^a*Steklov Mathematical Institute, Russian Academy of Sciences,
8 Gubkina Street, Moscow 119333, Russia*

Abstract. An integer random walk $\{S_i, i \geq 0\}$ with zero drift and finite variance σ^2 is investigated. For a random process that assigns, to a variable $u > 0$, the number of hits of the specified random walk into the state $\lfloor u\sigma\sqrt{n} \rfloor$ up to time n and is considered under the condition that $S_1 > 0, \dots, S_{n-1} > 0, S_n \leq 0$, a functional limit theorem concerning convergence in distribution of the process to the local time of a Brownian excursion is proved.

Keywords: random walk; conditional Brownian motion; local time of conditional Brownian motions; functional limit theorem.

Statement of the problem and the main result

Let $X_1, X_2, \dots$ be independent and identically distributed random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. We need the following assumptions about the distribution of the random variable X_1 .

Assumption 1. The random variable X_1 has an arithmetic distribution with the maximum span 1, and $\mathbf{P}(X_1 = 0) > 0$.

Instead of the above assumption, we can consider a more general one: the random variable X_1 has a lattice distribution, i. e. its possible values belong to the lattice $\{a + kh, k \in \mathbf{Z}\}$, where $h > 0$ and $a \in [0, h)$, the maximum of all possible h is 1, and the value of a corresponding to this h is 0.

Assumption 2. The following equalities hold:

$$\mathbf{E}X_1 = 0, \mathbf{E}X_1^2 = \sigma^2, 0 < \sigma^2 < +\infty.$$

Let us consider the random walk

$$S_0 = 0, S_i = \sum_{j=1}^i X_j, i \in \mathbf{N}.$$

Let T be the first hitting time of the semi-axis $(-\infty, 0]$ by this random walk:

$$T = \min\{i > 0 : S_i \leq 0\}.$$

We introduce for each $n \in \mathbf{N}$ the random process $Y_n = \{Y_n(t), t \in [0, 1]\}$, where

$$Y_n(t) = \frac{S_{\lfloor nt \rfloor}}{\sigma\sqrt{n}}, t \in [0, 1].$$

Let $W = \{W(t), t \geq 0\}$ be a standard Brownian motion and $W^{ex} = \{W^{ex}(t), t \in [0, 1]\}$ be a normalised Brownian excursion. Let us denote β and γ the nearest zeros of a Brownian trajectory to point 1 on the left and right, respectively. Set $\Delta = \gamma - \beta$. It is known (see, for example, [1, sect. 1]) that

$$\left\{ \frac{|W(\beta + t\Delta)|}{\sqrt{\Delta}}, t \in [0, 1] \right\} \stackrel{D}{=} W^{ex}, \quad (1)$$

where the symbol $\stackrel{D}{=}$ means equality in distribution.

Let the symbol $\xrightarrow{D}$ denote convergence in distribution of a sequence of random elements. The following functional limit theorem is established in paper [2]: if assumptions 1 and 2 are valid, then, as $n \rightarrow \infty$,

$$\{Y_n | T = n\} \xrightarrow{D} W^{ex} \quad (2)$$

(this means convergence in distribution in the space $D[0, 1]$ with the Skorokhod topology).

Let $\xi(k, n)$ denote the number of hits of the walk $\{S_i, 0 \leq i \leq n\}$ into the state $k \in \mathbf{Z}$:

$$\xi(k, n) = |\{i : 0 \leq i \leq n, S_i = k\}|.$$

We call the random variable $\xi(k, n)$ the local time of the random walk $\{S_i, i \geq 0\}$ at point k up to time n . Introduce for each $n \in \mathbf{N}$ the random process $Z_n = \{Z_n(u), u \in \mathbf{R}\}$, where

$$Z_n(u) = \frac{\sigma \xi\left(\left\lfloor \frac{u\sigma\sqrt{n}}{\sqrt{n}} \right\rfloor, n\right)}{\sqrt{n}}, u \in \mathbf{R}.$$

We are interested in the asymptotic behaviour of the process considered under the condition that $T = n$.

For $u \in \mathbf{R}$ and $t \geq 0$, the Brownian motion W has almost surely the local time $l(u, t)$ defined by the formula

$$l(u, t) = \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_0^t I_{[u-\varepsilon, u+\varepsilon]}(W(s)) ds,$$

where $I_{[u-\varepsilon, u+\varepsilon]}(\cdot)$ is the indicator of the set $[u-\varepsilon, u+\varepsilon]$. Furthermore, it is possible to choose variants of $l(u, t)$ for different t and u such that the function $l(u, t)$ will be almost surely continuous with respect to $(u, t) \in \mathbf{R} \times [0, +\infty)$. It follows from this fact and relation (1) that, firstly, for each $u \geq 0$ and $t \in [0, 1]$, there exists almost surely the local time of the Brownian excursion W^{ex} defined by the formula

$$l^{ex}(u, t) = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \int_0^t I_{[u, u+\varepsilon]}(W^{ex}(s)) ds,$$

and, secondly, it is possible to choose variants of $l^{ex}(u, t)$ for different t and u such that the function $l^{ex}(u, t)$ will be almost surely continuous with respect to $(u, t) \in [0, +\infty) \times [0, 1]$. This is assumed in what follows.

Theorem 1. *If assumptions 1 and 2 are valid, then, as $n \rightarrow \infty$,*

$$\{Z_n(u), u \geq 0 \mid T = n\} \xrightarrow{D} \{l^{ex}(u, 1), u \geq 0\}$$

(this means convergence in distribution in the space $D[0, v]$ with the Skorokhod topology for arbitrary $v > 0$).

As a corollary of theorem 1, the following results are obtained. Set for $x \in \mathbf{R}$

$$n(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

For each $u > 0$, let us introduce the distribution function

$$T_u^{ex}(x) = 1 - 2\sqrt{2\pi} \sum_{l=1}^{\infty} \sum_{j=0}^{l-1} C_{l-1}^j \frac{x^j}{j!} n^{(j+2)}(2lu + x), x \geq 0$$

($T_u^{ex}(x) = 0$ for $x < 0$). We also introduce the distribution function

$$F^{ex}(x) = 1 + 2 \sum_{l=1}^{\infty} e^{-2l^2 x^2} (1 - 4l^2 x^2), x > 0$$

($F^{ex}(x) = 0$ for $x \leq 0$). Note that the function $T_u^{ex}(x)$ is continuous for $x > 0$ and $T_u^{ex}(0) = F^{ex}(u)$.

Theorem 2. *If assumptions 1 and 2 are valid, then for every $u \geq 0$ and for all $x \geq 0$*

$$\lim_{n \rightarrow \infty} \mathbf{P}(Z_n(u) \leq x \mid T = n) = T_u^{ex}(x). \quad (3)$$

In particular,

$$\lim_{n \rightarrow \infty} \mathbf{P}(Z_n(u) = 0 \mid T = n) = F^{ex}(u). \quad (4)$$

Set

$$M_n = \sup_{u \geq 0} Z_n(u).$$

Theorem 3. *If assumptions 1 and 2 are valid, then for each $x \geq 0$*

$$\lim_{n \rightarrow \infty} \mathbf{P}(M_n \leq x \mid T = n) = F^{ex}\left(\frac{x}{2}\right).$$

This paper continues the author's series of papers, in which limit theorems on convergence to the local time of a Brownian meander [3], a stopped Brownian meander [4], a Brownian high jump [5], a Brownian

bridge and a reflected Brownian bridge [6] are established. Finally, we note that if the random walk $\{S_i, i \geq 0\}$ is symmetric simple, convergence of one-dimensional distributions in theorem 1 was established in paper [7].

Auxiliary results

It is known (see, for example, [8, chap. 12, § 7 and chap. 18, § 5]) that if assumption 2 is valid, then $0 < \mathbf{E}|S_T| < +\infty$ and, as $n \rightarrow \infty$,

$$\mathbf{P}(T > n) \sim \frac{c_1}{\sqrt{n}}, \quad (5)$$

where

$$c_1 = \frac{\sqrt{2}\mathbf{E}|S_T|}{\sqrt{\pi}\sigma}.$$

It was shown in paper [9] that if assumptions 1 and 2 are valid, then, as $n \rightarrow \infty$,

$$\mathbf{P}(T = n) \sim \frac{c_1}{2n^{\frac{3}{2}}}. \quad (6)$$

Let for $x < 0$

$$T(x) = \min\{n : S_n \in (-\infty, x]\}.$$

Let us introduce the distribution function G : for $y > 0$

$$G(y) = \frac{1}{\mathbf{E}|S_T|} (\mathbf{E}|S_T| - \mathbf{E}(|S_T| - y; |S_T| \geq y)).$$

The following result is established in paper [2, lemma 2]: if assumptions 1 and 2 are valid, then for all $x, y > 0$, as $n \rightarrow \infty$,

$$P(T(-\sigma\sqrt{n}x) = n, S_n + \sigma\sqrt{n}x > -y) = \frac{1}{n\sqrt{2\pi}} x e^{-\frac{x^2}{2}} G(y) (1 + o(1)), \quad (7)$$

moreover, for a fixed $y > 0$, relation (7) holds uniformly over $x \geq a$, where a is an arbitrary positive number.

Recall that $W = \{W(t), t \geq 0\}$ is the Brownian motion and β is the nearest zero of a Brownian trajectory to point 1 on the left, set $\Delta_1 = 1 - \beta$. Let $W^{me} = \{W^{me}(t), t \in [0, 1]\}$ be a Brownian meander. It is known (see, for example, [1, sect. 1]) that

$$\left\{ \frac{|W(\beta + t\Delta_1)|}{\sqrt{\Delta_1}}, t \in [0, 1] \right\} \stackrel{D}{=} W^{me}. \quad (8)$$

It follows from relation (8) that, firstly, for $u \geq 0$ and $t \in [0, 1]$, there is almost surely the local time of the Brownian meander W^{me} defined by the formula

$$l^{me}(u, t) = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \int_{[u, u+\varepsilon]} I(W^{me}(s)) ds,$$

and, secondly, it is possible to choose such variants of $l^{me}(u, t)$ for different u and t that the function $l^{me}(u, t)$ will be almost surely continuous with respect to $(u, t) \in [0, +\infty) \times [0, 1]$. This is assumed in what follows.

It is also known (see, for example, [1, sect. 4]) that, as $\delta \downarrow 0$,

$$\{W^{me}(t), t \in [0, 1] \mid W^{me}(1) \leq \delta\} \xrightarrow{D} \{W^{ex}(t), t \in [0, 1]\} \quad (9)$$

(convergence in distribution in the space $C[0, 1]$ with uniform convergence topology).

Recall Iglehart functional limit theorem (see [10; 11]): if assumption 2 is valid, then, as $n \rightarrow \infty$,

$$\{Y_n(t), t \in [0, 1] \mid T > n\} \xrightarrow{D} \{W^{me}(t), t \in [0, 1]\}. \quad (10)$$

It is established in paper [3] that if assumptions 1 and 2 are valid, then, as $n \rightarrow \infty$,

$$\{Z_n(u), u \geq 0 \mid T > n\} \xrightarrow{D} \{l^{me}(u, 1), u \geq 0\}$$

(convergence in distribution in the space $D[0, v]$ with the Skorokhod topology for arbitrary $v > 0$). This relation can be generalised: if assumptions 1 and 2 are valid, then for an arbitrary $t \in (0, 1)$, as $n \rightarrow \infty$,

$$\{Z_n(u, t), u \geq 0 \mid T > n\} \xrightarrow{D} \{l^{me}(u, t), u \geq 0\}, \quad (11)$$

where for $u \geq 0$

$$Z_n(u, t) = \frac{\sigma \xi\left(\left\lfloor \frac{u\sigma\sqrt{n}}{\sqrt{n}} \right\rfloor, \left\lfloor nt \right\rfloor\right)}{\sqrt{n}}.$$

Relation (11) is also valid if the variable u (in left part of relation (11)) is replaced by $u + \alpha_n(u)$, where $\alpha_n(u)$ tends, as $n \rightarrow \infty$, to 0 uniformly over $u \geq 0$. Note that relations (10) and (11) can be combined into one: if assumptions 1 and 2 are valid, then for an arbitrary $t \in [0, 1]$, as $n \rightarrow \infty$,

$$\{(Y_n(s), Z_n(u, t)), s \in [0, t], u \geq 0 \mid T > n\} \xrightarrow{D} \{(W^{me}(s), l^{me}(u, t)), s \in [0, t], u \geq 0\} \quad (12)$$

(convergence in distribution in the space $D[0, t] \times D[0, v]$ with the Skorokhod topology for arbitrary $v > 0$).

We introduce the following notation: N_s is a random variable having a normal distribution with parameters 0 and $s > 0$; $n_s(x)$ is distribution density of the random variable N_s , i. e.

$$n_s(x) = \frac{1}{\sqrt{2\pi s}} e^{-\frac{x^2}{2s}}, \quad x \in \mathbf{R};$$

$N_s(a, b)$ is distribution of the random variable N_s over intervals, i. e.

$$N_s(a, b) = \mathbf{P}(a < N_s \leq b), \quad a, b \in \mathbf{R}, \quad a \leq b$$

(herewith $N_0(0, b) = \frac{1}{2}$ for $b > 0$). We set for $s > 0$ and $x, y \in \mathbf{R}$

$$g_s(x, y) = n_s(x - y) - n_s(x + y),$$

$$h_s(x) = \frac{x}{\frac{3}{s^2}} \exp\left(-\frac{x^2}{2s}\right).$$

It is known (see, for example, [1, sect. 1]) that the Brownian meander W^{me} is an inhomogeneous Markov process with transition density $p^{me}(s, x; t, y)$, where for $0 < t \leq 1$ and $y > 0$

$$p^{me}(0, 0; t, y) = 2h_t(y)N_{1-t}(0, y), \quad (13)$$

and for $0 < s < t \leq 1$, $x > 0$ and $y > 0$

$$p^{me}(s, x; t, y) = g_{t-s}(x, y) \frac{N_{1-t}(0, y)}{N_{1-s}(0, x)}. \quad (14)$$

Set for $t \in [0, 1]$

$$W_t^{me} = \{W^{me}(s), s \in [0, t]\}, \quad W_t^{ex} = \{W^{ex}(s), s \in [0, t]\}$$

and for $t \in (0, 1)$

$$q_t(x) = \frac{h_t(x)}{\sqrt{2\pi}N_t(0, x)}.$$

Lemma 1. Let $t \in (0, 1)$ and f be a measurable non-negative and bounded mapping $D[0, t]$ with the Skorokhod topology into $\mathbf{R}$, and let this mapping be almost surely continuous on trajectories of the process W_t^{ex} . Then

$$\mathbf{E}f(W_t^{ex}) = \int_0^{+\infty} q_{1-t}(x) d_x \mathbf{E}(f(W_t^{me}); W^{me}(t) \leq x).$$

Proof. It follows from relation (9) by the continuous mapping theorem (see [12, theorem 5.1]) that, as $\delta \downarrow 0$,

$$\left\{ f(W^{me}) \mid W^{me}(1) \leq \delta \right\} \xrightarrow{D} f(W^{ex}).$$

From where, according to the dominated convergence theorem, we find that

$$\mathbf{E}f(W_t^{ex}) = \lim_{\delta \downarrow 0} \mathbf{E} \left(f(W_t^{me}) \mid W^{me}(1) \leq \delta \right). \quad (15)$$

Since the Brownian meander is an inhomogeneous Markov process, then

$$\mathbf{E} \left(f(W_t^{me}) \mid W^{me}(1) \leq \delta \right) = \int_0^{+\infty} \frac{\mathbf{P}^{(t,x)}(W^{me}(1) \leq \delta)}{\mathbf{P}(W^{me}(1) \leq \delta)} d_x \mathbf{E} \left(f(W_t^{me}); W^{me}(t) \leq x \right) \quad (16)$$

(the superscript of the symbol $\mathbf{P}^{(t,x)}$ means that the Brownian meander starts from point x at time t). Note that for $x > 0$

$$\mathbf{P}^{(t,x)}(W^{me}(1) \leq \delta) = \int_0^\delta p^{me}(t, x; 1, y) dy, \quad (17)$$

$$\mathbf{P}(W^{me}(1) \leq \delta) = \int_0^\delta p^{me}(0, 0; 1, y) dy. \quad (18)$$

Since, as $y \rightarrow 0$,

$$g_{1-t}(x, y) \sim \frac{2h_{1-t}(x)}{\sqrt{2\pi}} y, \quad h_1(y) \sim y,$$

then (see relations (13) and (14))

$$p^{me}(t, x; 1, y) \sim \frac{h_{1-t}(x)y}{\sqrt{2\pi}N_{1-t}(0, x)}, \quad p^{me}(0, 0; 1, y) \sim y. \quad (19)$$

We find from formula (19) that, as $\delta \downarrow 0$,

$$\int_0^\delta p^{me}(t, x; 1, y) dy \sim \frac{h_{1-t}(x)\delta^2}{2\sqrt{2\pi}N_{1-t}(0, x)}, \quad (20)$$

$$\int_0^\delta p^{me}(0, 0; 1, y) dy \sim \frac{\delta^2}{2}. \quad (21)$$

It follows from relations (17), (18), (20) and (21) that for $x > 0$

$$\lim_{\delta \downarrow 0} \frac{\mathbf{P}^{(t,x)}(W^{me}(1) \leq \delta)}{\mathbf{P}(W^{me}(1) \leq \delta)} = q_{1-t}(x). \quad (22)$$

We show that for all $x > 0$ and $y \in (0, 1]$

$$\frac{p^{me}(t, x; 1, y)}{y} \leq K(t), \quad (23)$$

where $K(t) > 0$ is independent of x, y . We find from relation (14) that

$$p^{me}(t, x; 1, y) = \frac{g_{1-t}(x, y)}{2N_{1-t}(0, x)} = \frac{n_{1-t}(y-x) - n_{1-t}(y+x)}{2 \int_0^x n_{1-t}(u) du}. \quad (24)$$

Since $1 - e^{-a} \leq a$ for $a \geq 0$, then

$$n_{1-t}(y-x) - n_{1-t}(y+x) = n_{1-t}(y-x) \left(1 - e^{\frac{-2xy}{1-t}} \right) \leq \frac{2xy}{1-t} n_{1-t}(y-x). \quad (25)$$

It follows from relations (24) and (25) that

$$\frac{p^{me}(t, x; 1, y)}{y} \leq \frac{xn_{1-t}(y-x)}{(1-t) \int_0^x n_{1-t}(u) du}. \quad (26)$$

Set $x_0 = \sqrt{2(1-t)\ln 2}$. It is not difficult to see that for $0 \leq u \leq x_0$

$$n_{1-t}(u) \geq \frac{1}{2\sqrt{2\pi(1-t)}}.$$

Therefore, if $0 < x \leq x_0$, then

$$\int_0^x n_{1-t}(u) du \geq \frac{x}{2\sqrt{2\pi(1-t)}}. \quad (27)$$

We get from relations (25)–(27) that for $0 < x \leq x_0$ and $y > 0$

$$\frac{p^{me}(t, x; 1, y)}{y} \leq \frac{2}{1-t} \exp\left(-\frac{(y-x)^2}{2(1-t)}\right) \leq \frac{2}{1-t},$$

i. e. relation (23) is valid for $0 < x \leq x_0$ and $y > 0$. In view of inequality (26), for $x \geq x_0$ and $y > 0$

$$\frac{p^{me}(t, x; 1, y)}{y} \leq \frac{xn_{1-t}(y-x)}{(1-t) \int_0^{x_0} n_{1-t}(u) du}. \quad (28)$$

It is easy to see that the numerator of the right-hand side of inequality (28) is bounded for $x > 0$ and $y \in (0, 1]$ by a positive constant depending on t . Hence, relation (23) is valid for $x \geq x_0$ and $y \in (0, 1]$. So, relation (23) is proved.

We have in view of relations (17) and (23) that for $\delta \in (0, 1]$ and any $x > 0$

$$\mathbf{P}^{(t, x)}(W^{me}(1) \leq \delta) \leq K(t) \frac{\delta^2}{2}. \quad (29)$$

From relation (29), taking into account formula (21), we find that there exists such $\delta_0 > 0$ that for any $\delta \in (0, \delta_0)$ and any $x > 0$

$$\frac{\mathbf{P}^{(t, x)}(W^{me}(1) \leq \delta)}{\mathbf{P}(W^{me}(1) \leq \delta)} \leq 2K(t). \quad (30)$$

According to the dominated convergence theorem and due to relations (22) and (30)

$$\lim_{\delta \downarrow 0} \int_0^{+\infty} \frac{\mathbf{P}^{(t, x)}(W^{me}(1) \leq \delta)}{\mathbf{P}(W^{me}(1) \leq \delta)} d_x \mathbf{E}(f(W_t^{me}); W^{me}(t) \leq x) = \int_0^{+\infty} q_{1-t}(x) d_x \mathbf{E}(f(W_t^{me}); W^{me}(t) \leq x).$$

Therefore (see relation (16)),

$$\lim_{\delta \downarrow 0} \mathbf{E}(f(W_t^{me}) \mid W^{me}(1) \leq \delta) = \int_0^{+\infty} q_{1-t}(x) d_x \mathbf{E}(f(W_t^{me}); W^{me}(t) \leq x).$$

From where, remembering formula (15), we come to the statement of lemma 1.

Let $0 \leq u < v$. Set for $t \in (0, 1]$ and $x \in D[0, t]$

$$f(x) = \int_0^t I_{[u, v]}(x(s)) ds.$$

This functional is measurable, and continuous for almost all trajectories of the process W^{me} (see [12, app. II]). Set

$$\mu^{me}(u, v, t) = f(W_t^{me}), \quad \mu^{ex}(u, v, t) = f(W_t^{ex}).$$

Note that $\mu^{me}(u, v, t)$ is the occupation time of the interval $[u, v]$ by the process W_t^{me} , and $\mu^{ex}(u, v, t)$ is the occupation time of the interval $[u, v]$ by the process W_t^{ex} . Moreover,

$$l^{me}(u, t) = \lim_{\varepsilon \downarrow 0} \frac{\mu^{me}(u, u + \varepsilon, t)}{\varepsilon}, \quad l^{ex}(u, t) = \lim_{\varepsilon \downarrow 0} \frac{\mu^{ex}(u, u + \varepsilon, t)}{\varepsilon}.$$

Let $m \in \mathbb{N}$, $\lambda_1, \dots, \lambda_m \geq 0$ and $u_1, \dots, u_m > 0$. Set

$$\Sigma_m^{me}(t) = \sum_{k=1}^m \lambda_k l^{me}(u_k, t),$$

$$\Sigma_m^{ex}(t) = \sum_{k=1}^m \lambda_k l^{ex}(u_k, t).$$

Lemma 2. For arbitrary $t \in (0, 1)$

$$\mathbf{E} e^{-\Sigma_m^{ex}(t)} = \int_0^{+\infty} q_{1-t}(x) d_x \mathbf{E} \left(e^{-\Sigma_m^{me}(t)}; W^{me}(t) \leq x \right).$$

Proof. By virtue of lemma 1 for $\varepsilon > 0$

$$\mathbf{E} \exp \left(- \sum_{k=1}^m \lambda_k \frac{\mu^{ex}(u_k, u_k + \varepsilon, t)}{\varepsilon} \right) = \int_0^{+\infty} q_{1-t}(x) d_x \mathbf{E} \left(\exp \left(- \sum_{k=1}^m \lambda_k \frac{\mu^{me}(u_k, u_k + \varepsilon)}{\varepsilon} \right); W^{me}(t) \leq x \right).$$

After integration in parts on the right side and passing to the limit, as $\varepsilon \rightarrow 0$, we obtain, according to the dominated convergence theorem, the statement of lemma 2.

For arbitrary $t \in (0, 1)$, we set

$$l_t^{me} = \{ l^{me}(u, t), u \geq 0 \}, \quad l_t^{ex} = \{ l^{ex}(u, t), u \geq 0 \}.$$

The following statement is of independent interest.

Corollary 1. For arbitrary $t \in (0, 1)$, $v > 0$ and any Borel set A of the space $D[0, v]$

$$\mathbf{P} \left(l_t^{ex} \in A \right) = \int_0^{+\infty} q_{1-t}(x) d_x \mathbf{P} \left(l_t^{me} \in A, W^{me}(t) \leq x \right). \quad (31)$$

Proof. By virtue of lemma 2, relation (31) is valid for cylindrical sets A of the space $D[0, v]$. Since both the left and right sides of relation (31) are probability measures, the required statement follows from Caratheodory theorem.

Proof of theorem 1

We establish in the beginning convergence of finite-dimensional distributions. Let us fix $m \in \mathbb{N}$ and $u_1, \dots, u_m > 0$. We show that, as $n \rightarrow \infty$,

$$(Z_n(u_1), \dots, Z_n(u_m) \mid T = n) \xrightarrow{D} (l^{ex}(u_1), \dots, l^{ex}(u_m)). \quad (32)$$

For an arbitrary $t \in (0, 1)$, we write the representation: for $u \geq 0$

$$Z_n(u) = Z_n(u, t) + \zeta_n(u, t), \quad (33)$$

where

$$\zeta_n(u, t) = \frac{\sigma \xi \left(\left\lfloor u \sigma \sqrt{n} \right\rfloor, \left\lfloor nt \right\rfloor + 1, n \right)}{\sqrt{n}},$$

and for $l \leq n$

$$\xi(j, l, n) = \left| \{ i \in \{l, \dots, n\} : S_i = j \} \right|.$$

First, we show that

$$\lim_{n \rightarrow \infty} \mathbf{E} \left(e^{-\Sigma_m(n, t)} \mid T = n \right) = \mathbf{E} e^{-\Sigma_m^{ex}(t)}, \quad (34)$$

where

$$\Sigma_m(n, t) = \sum_{k=1}^m \lambda_k Z_n(u_k, t)$$

and $\lambda_1, \dots, \lambda_m \geq 0$. For arbitrary $y, z > 0$

$$0 \leq \mathbf{E}\left(e^{-\Sigma_m(n, t)} \mid T = n\right) - E(n, y, z) \leq E_1(n, z) + E_2(n, y), \quad (35)$$

where

$$E(n, y, z) = \mathbf{E}\left(e^{-\Sigma_m(n, t)}; Y_n(t) > z, S_n > -y \mid T = n\right),$$

$$E_1(n, z) = \mathbf{E}\left(e^{-\Sigma_m(n, t)}; Y_n(t) \leq z \mid T = n\right),$$

$$E_2(n, y) = \mathbf{E}\left(e^{-\Sigma_m(n, t)}; S_n \leq -y \mid T = n\right).$$

Note that

$$0 \leq E_1(n, z) \leq \mathbf{P}(Y_n(t) \leq z \mid T = n). \quad (36)$$

Applying relation (2) to inequality (36), we find that

$$\limsup_{n \rightarrow \infty} E_1(n, z) \leq \mathbf{P}(W^{ex}(t) \leq z),$$

hence,

$$\lim_{z \rightarrow 0} \limsup_{n \rightarrow \infty} E_1(n, z) = 0. \quad (37)$$

It is shown in paper [2, relation (4.7)] that

$$\{S_n \mid T = n\} \xrightarrow{D} S, \quad (38)$$

where S is some random variable, and $\mathbf{P}(S > -y) = G(y)$ for $y > 0$. Since

$$0 \leq E_2(n, y) \leq \mathbf{P}(S_n \leq -y \mid T = n),$$

then (see formula (38))

$$\limsup_{n \rightarrow \infty} E_2(n, y) \leq \mathbf{P}(S \leq -y)$$

for any $y > 0$, with the exception of no more than a countable set. Hence,

$$\lim_{y \rightarrow \infty} \limsup_{n \rightarrow \infty} E_2(n, y) = 0. \quad (39)$$

Set $n(t) = n - \lfloor nt \rfloor$. By the Markov property of random walks

$$n^{\frac{3}{2}} \mathbf{E}\left(e^{-\Sigma_m(n, t)}; Y_n(t) > z, S_n > -y, T = n\right) = - \int_{(z, +\infty)} f_n(x) d_x \mathbf{E}\left(e^{-\Sigma_m(n, t)}; Y_n(t) > x \mid T > \lfloor nt \rfloor\right), \quad (40)$$

where for $x > 0$

$$f_n(x) = n^{\frac{3}{2}} \mathbf{P}(T > \lfloor nt \rfloor) \mathbf{P}(T(-\sigma\sqrt{nx}) = n(t), S_{n(t)} + \sigma\sqrt{nx} > -y).$$

We have by virtue of relations (5) and (7) that for arbitrary $x, y > 0$, as $n \rightarrow \infty$,

$$f_n(x) = \frac{c_1 G(y)}{\sqrt{2\pi t}} h_{1-t}(x) (1 + o(1)), \quad (41)$$

moreover, for a fixed y , this relation is performed uniformly over $x > z$.

Set

$$\tilde{\Sigma}_m^{me}(t) = \sum_{k=1}^m \lambda_k \tilde{l}^{me}(u_k, t),$$

where for $u > 0$

$$\tilde{l}^{me}(u, t) = \sqrt{t} l^{me}\left(\frac{u}{\sqrt{t}}, 1\right).$$

In view of formula (12)

$$\lim_{n \rightarrow \infty} \mathbf{E}\left(e^{-\Sigma_m(n,t)}; Y_n(t) > x \mid T > \lfloor nt \rfloor\right) = \mathbf{E}\left(e^{-\tilde{\Sigma}_m^{me}(t)}; \sqrt{t} W^{me}(1) > x\right). \quad (42)$$

It follows from relations (40)–(42) that

$$\lim_{n \rightarrow \infty} n^{\frac{3}{2}} \mathbf{E}\left(e^{-\Sigma_m(n,t)}; Y_n(t) > z, S_n > -y, T = n\right) = \frac{c_1 G(y)}{\sqrt{2\pi t}} \int_{(z, +\infty)} h_{1-t}(x) d_x \mathbf{E}\left(e^{-\tilde{\Sigma}_m^{me}(t)}; W^{me}(1) \leq \frac{x}{\sqrt{t}}\right). \quad (43)$$

Due to formulas (6) and (43)

$$\lim_{n \rightarrow \infty} E(n, y, z) = G(y) \sqrt{\frac{2}{\pi t}} \int_{(z, +\infty)} h_{1-t}(x) d_x \mathbf{E}\left(e^{-\tilde{\Sigma}_m^{me}(t)}; W^{me}(1) \leq \frac{x}{\sqrt{t}}\right).$$

Therefore,

$$\lim_{y \rightarrow \infty} \lim_{z \rightarrow 0} \lim_{n \rightarrow \infty} E(n, y, z) = \int_0^{+\infty} \sqrt{\frac{2}{\pi t}} h_{1-t}(x) d_x \mathbf{E}\left(e^{-\tilde{\Sigma}_m^{me}(t)}; W^{me}(1) \leq \frac{x}{\sqrt{t}}\right). \quad (44)$$

We find from relations (35), (37), (39) and (44) that

$$\lim_{n \rightarrow \infty} \mathbf{E}\left(e^{-\Sigma_m(n,t)} \mid T = n\right) = \sqrt{\frac{2}{\pi t}} \int_0^{+\infty} h_{1-t}(x) d_x \mathbf{E}\left(e^{-\tilde{\Sigma}_m^{me}(t)}; W^{me}(1) \leq \frac{x}{\sqrt{t}}\right). \quad (45)$$

Lemma 3. For arbitrary $t \in (0, 1)$ and any $x > 0$

$$\mathbf{E}\left(e^{-\Sigma_m^{me}(t)}; W^{me}(t) \leq x\right) = \frac{2}{\sqrt{t}} \int_0^x N_{1-t}(0, z) d_z \mathbf{E}\left(e^{-\tilde{\Sigma}_m^{me}(t)}; W^{me}(1) \leq \frac{z}{\sqrt{t}}\right).$$

Proof. In view of formula (12)

$$\mathbf{E}\left(e^{-\Sigma_m^{me}(t)}; W^{me}(t) \leq x\right) = \lim_{n \rightarrow \infty} \frac{\mathbf{E}\left(e^{-\Sigma_m(n,t)}; Y_n(t) \leq x, T > n\right)}{\mathbf{P}(T > n)}. \quad (46)$$

Set

$$L_n = \min_{0 \leq i \leq n} S_i.$$

By the Markov property of random walks

$$\mathbf{E}\left(e^{-\Sigma_m(n,t)}; Y_n(t) \leq x, T > n\right) = \int_0^x \mathbf{P}\left(\frac{L_n(t)}{\sigma\sqrt{n}} > -z\right) d_z \mathbf{E}\left(e^{-\Sigma_m(n,t)}; Y_n(t) \leq z, T > \lfloor nt \rfloor\right). \quad (47)$$

By virtue of the Donsker – Prokhorov invariance principle

$$\lim_{n \rightarrow \infty} \mathbf{P}\left(\frac{L_n(t)}{\sigma\sqrt{n}} > -z\right) = 2N_1\left(0, \frac{z}{\sqrt{1-t}}\right) = 2N_{1-t}(0, z), \quad (48)$$

and the first equality holds uniformly over $z \in (0, x]$. Due to formula (12)

$$\lim_{n \rightarrow \infty} \frac{\mathbf{E}\left(e^{-\Sigma_m(n,t)}; Y_n(t) \leq z, T > \lfloor nt \rfloor\right)}{\mathbf{P}(T > \lfloor nt \rfloor)} = \mathbf{E}\left(e^{-\tilde{\Sigma}_m^{me}(t)}; W^{me}(t) \leq \frac{z}{\sqrt{t}}\right). \quad (49)$$

It follows from relations (5) and (47)–(49) that

$$\lim_{n \rightarrow \infty} \frac{\mathbf{E}\left(e^{-\Sigma_m(n,t)}; Y_n(t) \leq x, T > n\right)}{\mathbf{P}(T > n)} = \frac{2}{\sqrt{t}} \int_0^x N_{1-t}(0, z) d_z \mathbf{E}\left(e^{-\tilde{\Sigma}_m^{me}(t)}; W^{me}(t) \leq \frac{z}{\sqrt{t}}\right). \quad (50)$$

Relations (46) and (50) imply the statement of lemma 3.

Let us continue the proof of theorem 1. We show that

$$\sqrt{\frac{2}{\pi t}} \int_0^{+\infty} h_{1-t}(x) d_x \mathbf{E}\left(e^{-\tilde{\Sigma}_m^{me}(t)}; W^{me}(1) \leq \frac{x}{\sqrt{t}}\right) = \mathbf{E}e^{-\Sigma_m^{ex}(t)}. \quad (51)$$

By lemma 2

$$\int_0^{+\infty} q_{1-t}(x) d_x \mathbf{E}\left(e^{-\Sigma_m^{me}(t)}; W^{me}(t) \leq x\right) = \mathbf{E}e^{-\Sigma_m^{ex}(t)}. \quad (52)$$

By lemma 3

$$\begin{aligned} \int_0^{+\infty} q_{1-t}(x) d_x \mathbf{E}\left(e^{-\Sigma_m^{me}(t)}; W^{me}(t) \leq x\right) &= \frac{2}{\sqrt{t}} \int_0^{+\infty} q_{1-t}(x) N_{1-t}(0, x) d_x \mathbf{E}\left(e^{-\tilde{\Sigma}_m^{me}(t)}; W^{me}(1) \leq \frac{x}{\sqrt{t}}\right) = \\ &= \sqrt{\frac{2}{\pi t}} \int_0^{+\infty} h_{1-t}(x) d_x \mathbf{E}\left(e^{-\tilde{\Sigma}_m^{me}(t)}; W^{me}(1) \leq \frac{x}{\sqrt{t}}\right). \end{aligned} \quad (53)$$

We obtain from formulas (52) and (53) the required relation (51), and from formulas (45) and (51) we have the required relation (34).

By the continuity theorem for characteristic functions, we find from relation (34) that, as $n \rightarrow \infty$,

$$(Z_n(u_1, t), \dots, Z_n(u_m, t) \mid T = n) \xrightarrow{D} (l^{ex}(u_1, t), \dots, l^{ex}(u_m, t)). \quad (54)$$

Since the function $l^{ex}(u, t)$ is continuous with respect t with probability 1, then almost surely, as $t \rightarrow 1$,

$$(l^{ex}(u_1, t), \dots, l^{ex}(u_m, t)) \rightarrow (l^{ex}(u_1, 1), \dots, l^{ex}(u_m, 1)). \quad (55)$$

We show that for an arbitrary $u > 0$

$$\lim_{t \rightarrow 1} \limsup_{n \rightarrow \infty} \mathbf{P}(\zeta_n(u, t) > 0 \mid T = n) = 0. \quad (56)$$

If $\sup_{t \leq s \leq 1} Y_n(s) < u$, then $\zeta_n(u, t) = 0$. Therefore,

$$\mathbf{P}(\zeta_n(u, t) > 0 \mid T = n) = \mathbf{P}\left(\zeta_n(u, t) > 0, \sup_{t \leq s \leq 1} Y_n(s) \geq u \mid T = n\right). \quad (57)$$

Further,

$$\mathbf{P}\left(\zeta_n(u, t) > 0, \sup_{t \leq s \leq 1} Y_n(s) \geq u \mid T = n\right) \leq \mathbf{P}\left(\sup_{t \leq s \leq 1} Y_n(s) \geq u \mid T = n\right). \quad (58)$$

By virtue of relation (2)

$$\lim_{n \rightarrow \infty} \mathbf{P}\left(\sup_{t \leq s \leq 1} Y_n(s) \geq u \mid T = n\right) = \mathbf{P}\left(\sup_{t \leq s \leq 1} W^{ex}(s) \geq u\right). \quad (59)$$

It follows from formulas (57)–(59) that

$$\limsup_{n \rightarrow \infty} \mathbf{P}(\zeta_n(u, t) > 0 \mid T = n) \leq \mathbf{P}\left(\sup_{t \leq s \leq 1} W^{ex}(s) \geq u\right). \quad (60)$$

Since almost surely $\sup_{t \leq s \leq 1} W^{ex}(s) \rightarrow 0$, as $t \rightarrow 1$, then

$$\lim_{t \rightarrow 1} \mathbf{P}\left(\sup_{t \leq s \leq 1} W^{ex}(s) \geq u\right) = 0. \quad (61)$$

We obtain from formulas (60) and (61) the required relation (56).

In view of relation (56)

$$\lim_{t \rightarrow 1} \limsup_{n \rightarrow \infty} \mathbf{P} \left(\max_{1 \leq k \leq m} \zeta_n(u_k, t) > 0 \mid T = n \right) = 0. \quad (62)$$

We find from formulas (33) and (62) that

$$\lim_{t \rightarrow 1} \limsup_{n \rightarrow \infty} \mathbf{P} \left(\max_{1 \leq k \leq m} |Z_n(u_k) - Z_n(u_k, t)| > 0 \mid T = n \right) = 0. \quad (63)$$

By applying formulas (54), (55), (63) and theorem 4.2 from paper [12], we obtain the required relation (32).

Let us fix $v > 0$. We introduce a continuity module for $x \in D[0, v]$: for $\delta > 0$

$$w_x(\delta) = \sup_{u_1, u_2 \in [0, v]: |u_1 - u_2| \leq \delta} |x(u_1) - x(u_2)|.$$

We show that for an arbitrary $\varepsilon > 0$

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbf{P} \left(w_{Z_n}(\delta) \geq \varepsilon \mid T = n \right) = 0. \quad (64)$$

Note (see formula (33)) that for $t \in (0, 1)$

$$w_{Z_n}(\delta) \leq w_{Z_n(\cdot, t)}(\delta) + w_{\zeta_n(\cdot, t)}(\delta). \quad (65)$$

Due to relation (65)

$$\mathbf{P} \left(w_{Z_n}(\delta) \geq \varepsilon \mid T = n \right) \leq \mathbf{P} \left(w_{Z_n(\cdot, t)}(\delta) \geq \frac{\varepsilon}{2} \mid T = n \right) + \mathbf{P} \left(w_{\zeta_n(\cdot, t)}(\delta) \geq \frac{\varepsilon}{2} \mid T = n \right). \quad (66)$$

First, we show that

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbf{P} \left(w_{Z_n(\cdot, t)}(\delta) \geq \frac{\varepsilon}{2} \mid T = n \right) = 0. \quad (67)$$

Note that for arbitrary $y, z > 0$

$$0 \leq \mathbf{P} \left(w_{Z_n(\cdot, t)}(\delta) \geq \frac{\varepsilon}{2} \mid T = n \right) - P(n, y, z) \leq P_1(n, z) + P_2(n, y), \quad (68)$$

where

$$P(n, y, z) = \mathbf{P} \left(w_{Z_n(\cdot, t)}(\delta) \geq \frac{\varepsilon}{2}, Y_n(t) > z, S_n > -y \mid T = n \right),$$

$$P_1(n, z) = \mathbf{P} \left(w_{Z_n(\cdot, t)}(\delta) \geq \frac{\varepsilon}{2}, Y_n(t) \leq z \mid T = n \right),$$

$$P_2(n, y) = \mathbf{P} \left(w_{Z_n(\cdot, t)}(\delta) \geq \frac{\varepsilon}{2}, S_n \leq -y \mid T = n \right).$$

Since

$$P_1(n, z) \leq \mathbf{P} \left(Y_n(t) \leq z \mid T = n \right),$$

then (see relation (2))

$$\limsup_{n \rightarrow \infty} P_1(n, z) \leq \mathbf{P} \left(W^{ex}(t) \leq z \right).$$

Therefore,

$$\lim_{z \rightarrow 0} \limsup_{n \rightarrow \infty} P_1(n, z) = 0. \quad (69)$$

Since

$$P_2(n, y) \leq \mathbf{P} \left(S_n \leq -y \mid T = n \right),$$

then (see formula (38))

$$\limsup_{n \rightarrow \infty} P_2(n, y) \leq \mathbf{P}(S \leq -y)$$

for any $y > 0$, with the exception of no more than a countable set. Therefore,

$$\lim_{y \rightarrow \infty} \limsup_{n \rightarrow \infty} P_2(n, y) = 0. \quad (70)$$

By the Markov property of random walks

$$n^{\frac{3}{2}} \mathbf{P} \left(w_{Z_n(\cdot, t)}(\delta) \geq \frac{\varepsilon}{2}, Y_n(t) > z, S_n > -y, T = n \right) = - \int_{(z, +\infty)} f_n(x) d_x \mathbf{P} \left(w_{Z_n(\cdot, t)}(\delta) \geq \frac{\varepsilon}{2}, Y_n(t) > x \mid T > \lfloor nt \rfloor \right), \quad (71)$$

where $f_n(x)$ is the same function as in relation (40). We have in view of formula (41) that for a fixed $y > 0$, all $x > z$ and all sufficiently large n

$$f_n(x) \leq \frac{2c_1 G(y)}{\sqrt{2\pi t}} h_{1-t}(x) \leq K(t), \quad (72)$$

where $K(t) > 0$ is independent of n, x . We find from formulas (71) and (72) that for a fixed $y > 0$ and all sufficiently large n

$$\begin{aligned} n^{\frac{3}{2}} \mathbf{P} \left(w_{Z_n(\cdot, t)}(\delta) \geq \frac{\varepsilon}{2}, Y_n(t) > z, S_n > -y, T = n \right) &\leq \\ &\leq K(t) \mathbf{P} \left(w_{Z_n(\cdot, t)}(\delta) \geq \frac{\varepsilon}{2}, Y_n(t) > z \mid T > \lfloor nt \rfloor \right) \leq K(t) \mathbf{P} \left(w_{Z_n(\cdot, t)}(\delta) \geq \frac{\varepsilon}{2} \mid T > \lfloor nt \rfloor \right). \end{aligned} \quad (73)$$

It follows from relation (11) that

$$\lim_{\delta \rightarrow 0} \lim_{n \rightarrow \infty} \mathbf{P} \left(w_{Z_n(\cdot, t)}(\delta) \geq \frac{\varepsilon}{2} \mid T > \lfloor nt \rfloor \right) = 0. \quad (74)$$

We get from formulas (73) and (74) that

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} n^{\frac{3}{2}} \mathbf{P} \left(w_{Z_n(\cdot, t)}(\delta) \geq \frac{\varepsilon}{2}, Y_n(t) > z, S_n > -y, T = n \right) = 0$$

and, hence (see relation (6)),

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} P(n, y, z) = 0. \quad (75)$$

We obtain from formulas (68)–(70) and (75) the required relation (67).

Now let us show that

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbf{P} \left(w_{\zeta_n(\cdot, t)}(\delta) \geq \frac{\varepsilon}{2} \mid T = n \right). \quad (76)$$

Note that for an arbitrary $L > 0$

$$0 \leq \mathbf{P} \left(w_{\zeta_n(\cdot, t)}(\delta) \geq \frac{\varepsilon}{2} \mid T = n \right) - \sum_{l=0}^{\lfloor L \rfloor} q(n, l) \leq Q(n, L), \quad (77)$$

where

$$Q(n, L) = \mathbf{P}(S_n < -L \mid T = n)$$

and for $l \in \{0, 1, 2, \dots\}$

$$q(n, l) = \mathbf{P} \left(w_{\zeta_n(\cdot, t)}(\delta) \geq \frac{\varepsilon}{2}, S_n = -l \mid T = n \right).$$

By virtue of formula (38)

$$\lim_{n \rightarrow \infty} Q(n, L) = \mathbf{P}(S < -L)$$

for any L , with the exception of no more than a countable set, hence,

$$\lim_{L \rightarrow \infty} \lim_{n \rightarrow \infty} Q(n, L) = 0. \quad (78)$$

We show that for a fixed $l \in \{0, 1, 2, \dots\}$

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} q(n, l) = 0. \quad (79)$$

Set

$$\tilde{S}_0 = 0, \tilde{S}_1 = S_{n-1} - S_n, \tilde{S}_2 = S_{n-2} - S_n, \dots, \tilde{S}_n = S_0 - S_n = -S_n.$$

Note that

$$(\tilde{S}_1, \dots, \tilde{S}_n) \stackrel{D}{=} -(S_1, \dots, S_n).$$

Let $\tilde{\xi}(k, m)$, where $m \in \{0, 1, \dots, n\}$, be the number of hits of the random walk $\{\tilde{S}_i, 0 \leq i \leq m\}$ in the state $k \in \mathbf{Z}$, i. e.

$$\tilde{\xi}(k, m) = |\{i : 0 \leq i \leq m, \tilde{S}_i = k\}|.$$

If $S_n = -l$, then for $m \in \{0, 1, \dots, n\}$

$$\begin{aligned} \tilde{\xi}(k, m) &= |\{i : 0 \leq i \leq m, S_{n-i} - S_n = k\}| = |\{i : 0 \leq i \leq m, S_{n-i} = k - l\}| = \\ &= |\{i : n - m \leq i \leq n, S_i = k - l\}| = \xi(k - l, n - m, n). \end{aligned} \quad (80)$$

Let $\tilde{T}$ mean the same for the walk $\{\tilde{S}_i\}$ as T means for the walk $\{S_i\}$. Then

$$\begin{aligned} \{S_n = -l, T = n\} &= \{S_1 > 0, \dots, S_{n-1} > 0, S_n = -l\} = \{\tilde{S}_{n-1} - \tilde{S}_n > 0, \dots, \tilde{S}_1 - \tilde{S}_n > 0, \tilde{S}_n = l\} = \\ &= \{\tilde{S}_1 > l, \dots, \tilde{S}_{n-1} > l, \tilde{S}_n = l\} \subset \{\tilde{S}_1 > 0, \dots, \tilde{S}_{n-1} > 0, \tilde{S}_n = l\} = \{\tilde{S}_n = l, \tilde{T} > n\}. \end{aligned} \quad (81)$$

Set for $u \in \mathbf{R}$

$$\tilde{Z}_n(u, t) = \frac{\sigma \tilde{\xi}(\lfloor u\sigma\sqrt{n} \rfloor + l, \lfloor nt \rfloor)}{\sqrt{n}}.$$

If $S_n = -l$, then by virtue of relation (80)

$$\zeta_n(u, t) = \frac{\sigma \tilde{\xi}(\lfloor u\sigma\sqrt{n} \rfloor + l, n - \lfloor nt \rfloor - 1)}{\sqrt{n}} = \frac{\sigma \tilde{\xi}(\lfloor u\sigma\sqrt{n} \rfloor + l, \lfloor n(1 - t_n) \rfloor)}{\sqrt{n}} = \tilde{Z}_n(u, 1 - t_n), \quad (82)$$

where t_n is found from the equality $n - \lfloor nt \rfloor - 1 = \lfloor n(1 - t_n) \rfloor$. Since $|t_n - t| \leq \frac{1}{n}$, then

$$\left| \tilde{\xi}(\lfloor u\sigma\sqrt{n} \rfloor + l, \lfloor n(1 - t_n) \rfloor) - \tilde{\xi}(\lfloor u\sigma\sqrt{n} \rfloor + l, \lfloor n(1 - t) \rfloor) \right| \leq 1$$

and, hence,

$$|\tilde{Z}_n(u, 1 - t_n) - \tilde{Z}_n(u, 1 - t)| \leq \frac{\sigma}{\sqrt{n}}. \quad (83)$$

It follows from inequality (83) that

$$|w_{\tilde{Z}_n(\cdot, 1 - t_n)}(\delta) - w_{\tilde{Z}_n(\cdot, 1 - t)}(\delta)| \leq \frac{2\sigma}{\sqrt{n}}. \quad (84)$$

If $S_n = -l$, then by virtue of relations (82) and (84)

$$|w_{\zeta_n(\cdot, t)}(\delta) - w_{\tilde{Z}_n(\cdot, 1 - t)}(\delta)| \leq \frac{2\sigma}{\sqrt{n}}. \quad (85)$$

To prove relation (79), it is enough to establish (see formula (6)) that

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} n^{\frac{3}{2}} \mathbf{P} \left(w_{\zeta_n(\cdot, t)}(\delta) \geq \frac{\varepsilon}{2}, S_n = -l, T = n \right) = 0. \quad (86)$$

In turn, to prove relation (86), it is enough to establish (see formulas (81), (82) and (85)) that

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} n^{\frac{3}{2}} \mathbf{P} \left(w_{\tilde{Z}_n(u, 1-t)}(\delta) \geq \frac{\varepsilon}{2}, \tilde{S}_n = l, \tilde{T} > n \right) = 0 \quad (87)$$

or, equivalently,

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} n^{\frac{3}{2}} \mathbf{P} \left(w_{Z_n(u, 1-t)}(\delta) \geq \frac{\varepsilon}{2}, S_n = l, T > n \right) = 0. \quad (88)$$

It is not difficult to show (see the proof of relation (4.20) in paper [2]) that there are constants $K > 0$ and $L \in \mathbf{N}$ such that

$$\mathbf{P} \left(w_{Z_n(\cdot, 1-t)}(\delta) \geq \frac{\varepsilon}{2}, S_n = l, T > n \right) \leq K \sum_{i=1}^L \mathbf{P} \left(w_{Z_n(\cdot, 1-t)}(\delta) \geq \frac{\varepsilon}{2}, T = n + i \right). \quad (89)$$

It follows from formula (6) and the proved relation (67) that for each $i \in \{1, \dots, l\}$

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} n^{\frac{3}{2}} \mathbf{P} \left(w_{Z_{n+i}(\cdot, 1-t)}(\delta) \geq \frac{\varepsilon}{2}, T = n + i \right) = 0. \quad (90)$$

We show that for each $i \in \{1, \dots, l\}$

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} n^{\frac{3}{2}} \mathbf{P} \left(w_{Z_n(\cdot, 1-t)}(\delta) \geq \frac{\varepsilon}{2}, T = n + i \right) = 0. \quad (91)$$

Set

$$Z_{n+i}^*(u, 1-t) = \frac{\sigma \xi \left(\left\lfloor u \sigma \sqrt{n+i} \right\rfloor, \left\lfloor n(1-t) \right\rfloor \right)}{\sqrt{n+i}}.$$

Since $\left\lfloor (n+i)t \right\rfloor - \left\lfloor nt \right\rfloor \leq i+1$, then

$$Z_{n+i}(u, 1-t) - Z_{n+i}^*(u, 1-t) \leq \sigma \frac{i+1}{\sqrt{n+i}}$$

and, hence,

$$\left| w_{Z_{n+i}(\cdot, 1-t)}(\delta) - w_{Z_{n+i}^*(\cdot, 1-t)}(\delta) \right| \leq \frac{2\sigma(i+1)}{\sqrt{n+i}}. \quad (92)$$

Further,

$$w_{Z_{n+i}^*(\cdot, 1-t)}(\delta) = \sqrt{\frac{n}{n+i}} w_{Z_n(\cdot, 1-t)} \left(\delta \sqrt{\frac{n+i}{n}} \right). \quad (93)$$

By virtue of formulas (92) and (93)

$$\left| w_{Z_{n+i}(\cdot, 1-t)}(\delta) - \sqrt{\frac{n}{n+i}} w_{Z_n(\cdot, 1-t)} \left(\delta \sqrt{\frac{n+i}{n}} \right) \right| \leq \frac{2\sigma(i+1)}{\sqrt{n+i}}$$

and, hence,

$$\left| \sqrt{\frac{n+i}{n}} w_{Z_{n+i}(\cdot, 1-t)} \left(\delta \sqrt{\frac{n}{n+i}} \right) - w_{Z_n(\cdot, 1-t)}(\delta) \right| \leq \frac{2\sigma(i+1)}{\sqrt{n}}.$$

Thus,

$$w_{Z_n(\cdot, 1-t)}(\delta) \leq \sqrt{\frac{n+i}{n}} w_{Z_{n+i}(\cdot, 1-t)} \left(\delta \sqrt{\frac{n}{n+i}} \right) + \frac{2\sigma(i+1)}{\sqrt{n}}$$

and, hence, for sufficiently large n

$$w_{Z_n(\cdot, 1-t)}(\delta) \leq 2w_{Z_{n+i}(\cdot, 1-t)}(\delta) + \frac{2\sigma(i+1)}{\sqrt{n}}. \quad (94)$$

It follows from formulas (90) and (94) the required relation (91).

From formulas (89) and (91) we have relation (88) and, hence, relations (87) and (86). In turn, formula (86) implies the required relation (79). Further, from formulas (77)–(79) we find the required relation (76). Finally, from formulas (66), (67) and (76) we find the required relation (64). From relations (32) and (64) according to theorem 15.5 from book [12] we obtain the statement of theorem 1.

Proof of theorem 2

By theorem 1, as $n \rightarrow \infty$,

$$\{Z_n(u) \mid T=n\} \xrightarrow{D} l^{ex}(u, 1). \quad (95)$$

Set for $x \in \mathbf{R}$

$$F_u(x) = \mathbf{P}(l^{ex}(u, 1) \leq x).$$

It follows from relation (95) that

$$\lim_{n \rightarrow \infty} \mathbf{P}(Z_n(u) \leq x \mid T=n) = F_u(x) \quad (96)$$

for all x , which are points of continuity of the function $F_u(\cdot)$. It is proved in paper [7, relations (170) and (171)] that for any $u > 0$ and for all $x \geq 0$

$$F_u(x) = T_u^{ex}(x). \quad (97)$$

The function $T_u^{ex}(x)$ is continuous for all $x > 0$ and, hence, relation (96) is valid for all $x > 0$. Formulas (96) and (97) imply relation (3) for $x > 0$.

Let us prove the relation (4). We have due to formula (33) that for $t \in (0, 1)$

$$\left| \mathbf{P}(Z_n(u) = 0 \mid T=n) - P_1(n, t, u) \right| \leq 1 - P_2(n, t, u), \quad (98)$$

where

$$\begin{aligned} P_1(n, t, u) &= \mathbf{P}(Z_n(u, t) = 0 \mid T=n), \\ P_2(n, t, u) &= \mathbf{P}(\zeta_n(u, t) = 0 \mid T=n). \end{aligned}$$

Similarly to relation (34), it can be shown that

$$\lim_{n \rightarrow \infty} P_1(n, t, u) = \mathbf{P}(l_t^{ex}(u) = 0). \quad (99)$$

Note that

$$\mathbf{P}(l_t^{ex}(u) = 0) = \mathbf{P}\left(\sup_{0 \leq s \leq t} W^{ex}(s) \leq u\right). \quad (100)$$

Trajectories of the Brownian excursion W^{ex} are almost surely continuous, and therefore,

$$\lim_{t \rightarrow 1} \mathbf{P}\left(\sup_{0 \leq s \leq t} W^{ex}(s) \leq u\right) = \mathbf{P}\left(\sup_{0 \leq s \leq 1} W^{ex}(s) \leq u\right). \quad (101)$$

It is known (see [1, sect. 6]) that

$$\mathbf{P}\left(\sup_{0 \leq s \leq 1} W^{ex}(s) \leq u\right) = F^{ex}(u). \quad (102)$$

It follows from formulas (100)–(102) that

$$\lim_{t \rightarrow 1} \mathbf{P}(l_t^{ex}(u) = 0) = F^{ex}(u). \quad (103)$$

Applying relation (103) to relation (99), we find that

$$\lim_{t \rightarrow 1} \lim_{n \rightarrow \infty} P_1(n, t, u) = F^{ex}(u). \quad (104)$$

It follows from relation (56) that

$$\lim_{t \rightarrow 1} \liminf_{n \rightarrow \infty} P_2(n, t, u) = 1. \quad (105)$$

From formulas (98), (104) and (105) we obtain the required relation (4). Theorem 2 is proved.

Proof of theorem 3

By virtue of theorem 1, as $n \rightarrow \infty$,

$$\sup_{u \geq 0} Z_n(u) \xrightarrow{D} \sup_{u \geq 0} l^{ex}(u, 1).$$

Hence,

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\sup_{u \geq 0} Z_n(u) \leq x \right) = \mathbf{P} \left(\sup_{u \geq 0} l^{ex}(u, 1) \leq x \right) \quad (106)$$

for all $x \geq 0$, which are points of continuity of the right-hand side. It is known (see, for example, [13, sect. 4.2]) that

$$\sup_{u \geq 0} l^{ex}(u, 1) \stackrel{D}{=} 2 \sup_{s \in [0, 1]} W^{ex}(s). \quad (107)$$

We find from formulas (102) and (107) that for $x \geq 0$

$$\mathbf{P} \left(\sup_{u \geq 0} l^{ex}(u, 1) \leq x \right) = \mathbf{P} \left(\sup_{s \in [0, 1]} W^{ex}(s) \leq \frac{x}{2} \right) = F^{ex} \left(\frac{x}{2} \right). \quad (108)$$

Since the right-hand side of relation (108) is continuous, then relation (106) holds for all $x \geq 0$. We obtain from relations (106) and (108) the statement of theorem 3.

References

1. Durrett RT, Iglehart DL, Miller DR. Weak convergence to Brownian meander and Brownian excursion. *The Annals of Probability*. 1977;5(1):117–129. DOI: 10.1214/aop/1176995895.
2. Afanasyev VI. Local invariance principle for a random walk with zero drift. *Journal of Mathematical Sciences*. 2022;266(6): 850–868. DOI: 10.1007/s10958-022-06145-8.
3. Afanasyev VI. Convergence to the local time of Brownian meander. *Discrete Mathematics and Applications*. 2019;29(3): 149–158. DOI: 10.1515/dma-2019-0014.
4. Afanasyev VI. Functional limit theorem for the local time of stopped random walk. *Discrete Mathematics and Applications*. 2020;30(3):147–157. DOI: 10.1515/dma-2020-0014.
5. Afanasyev VI. On the local time of a stopped random walk attaining a high level. *Proceedings of the Steklov Institute of Mathematics*. 2022;316:5–25. DOI: 10.1134/S0081543822010035.
6. Afanasyev VI. Limit theorem on convergence to the local time of a Brownian bridge. *Mathematical Notes*. 2024;116(5–6): 875–891. DOI: 10.1134/S0001434624110014.
7. Takács L. Brownian local times. *Journal of Applied Mathematics and Stochastic Analysis*. 1995;8(3):209–232. DOI: 10.1155/S1048953395000207.
8. Feller W. *An introduction to probability theory and its applications. Volume 2*. 2nd edition. New York: John Wiley and Sons; 1971. XXIV, 669 p. (Wiley series in probability and mathematical statistics).
9. Vatutin VA, Wachtel V. Local probabilities for random walks conditioned to stay positive. *Probability Theory and Related Fields*. 2009;143(1–2):177–217. DOI: 10.1007/s00440-007-0124-8.
10. Iglehart DL. Functional central limit theorems for random walks conditioned to stay positive. *The Annals of Probability*. 1974; 2(4):608–619. DOI: 10.1214/aop/1176996607.
11. Bolthausen E. On a functional central limit theorem for random walks conditioned to stay positive. *The Annals of Probability*. 1976;4(3):480–485. DOI: 10.1214/aop/1176996098.
12. Billingsley P. *Convergence of probability measures*. New York: John Wiley and Sons; 1968. XII, 253 p. (Wiley series in probability and mathematical statistics).
13. Drmota M. *Random trees: an interplay between combinatorics and probability*. Vienna: Springer-Verlag; 2009. XVII, 458 p. DOI: 10.1007/978-3-211-75357-6.

Received 24.12.2025 / revised 19.01.2026 / accepted 29.01.2026.

ОЦЕНИВАНИЕ ПАРАМЕТРОВ ПУАССОНОВСКОГО ПРОЦЕССА ВТОРОГО ПОРЯДКА, УПРАВЛЯЕМОГО МАРКОВСКИМ ПРОЦЕССОМ, С ИСПОЛЬЗОВАНИЕМ ПРОЦЕДУРЫ ОБНАРУЖЕНИЯ РАЗЛАДКИ

С. Э. ВОРОБЕЙЧИКОВ¹⁾, Ю. Б. БУРКАТОВСКАЯ^{1), 2)}

¹⁾Томский государственный университет, пр. Ленина, 36, 634050, г. Томск, Россия

²⁾Томский политехнический университет, пр. Ленина, 30, 634050, г. Томск, Россия

Аннотация. Рассматривается проблема оценки параметров пуассоновского процесса второго порядка, управляемого марковским процессом. Разрабатывается новый двухэтапный алгоритм их оценки. На первом этапе строятся начальные оценки параметров интенсивности с использованием метода моментов и метода максимального правдоподобия. На втором этапе эти оценки применяются в алгоритме кумулятивных сумм обнаружения разладки для определения временных интервалов, при которых интенсивность процесса остается постоянной; затем строятся оценки всех параметров. Разработанный алгоритм сравнивается с известным алгоритмом для нахождения оценок максимального правдоподобия (ЕМ-алгоритмом). Предложенный метод не является итерационным, что обеспечивает довольно низкую вычислительную сложность.

Ключевые слова: пуассоновский процесс, управляемый марковским процессом; кумулятивные суммы; метод моментов; метод максимального правдоподобия.

Благодарность. Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (грант № 24-11-00191 на имя С. Э. Воробейчикова).

Образец цитирования:

Воробейчиков СЭ, Буркатовская ЮБ. Оценивание параметров пуассоновского процесса второго порядка, управляемого марковским процессом, с использованием процедуры обнаружения разладки. *Журнал Белорусского государственного университета. Математика. Информатика.* 2026;1:53–60. (на англ.).
EDN: FJKYPG

For citation:

Vorobeychikov SE, Burkatovskaya YuB. Estimation of the parameters of a second order Markov modulated Poisson process using change-point detection procedure. *Journal of the Belarusian State University. Mathematics and Informatics.* 2026;1: 53–60.
EDN: FJKYPG

Авторы:

Сергей Эрикович Воробейчиков – доктор физико-математических наук, доцент; профессор кафедры системного анализа и математического моделирования Института прикладной математики и компьютерных наук.

Юлия Борисовна Буркатовская – кандидат физико-математических наук, доцент; доцент кафедры системного анализа и математического моделирования Института прикладной математики и компьютерных наук¹⁾, доцент отделения информационных технологий Инженерной школы информационных технологий и робототехники²⁾.

Authors:

Sergey E. Vorobeychikov, doctor of science (physics and mathematics), docent; professor at the department of system analysis and mathematical modelling, Institute of Applied Mathematics and Computer Science.

sergeivorobeichikov@yandex.ru

<https://orcid.org/0000-0002-3928-427X>

Yulia B. Burkatovskaya, PhD (physics and mathematics), docent; associate professor at the department of system analysis and mathematical modelling, Institute of Applied Mathematics and Computer Science^a, and associate professor at the division for information technology, School of Computer Science and Robotics^b.

tracey@tpu.ru

<https://orcid.org/0000-0002-0720-7560>



ESTIMATION OF THE PARAMETERS OF A SECOND ORDER MARKOV MODULATED POISSON PROCESS USING CHANGE-POINT DETECTION PROCEDURE

S. E. VOROBAYCHIKOV^a, Yu. B. BURKATOVSKAYA^{a, b}

^aNational Research Tomsk State University, 36 Lenina Avenue, Tomsk 634050, Russia

^bNational Research Tomsk Polytechnic University, 30 Lenina Avenue, Tomsk 634050, Russia

Corresponding author: Yu. B. Burkatovskaya (tracey@tpu.ru)

Abstract. The problem of estimating the parameters of a second order Markov modulated Poisson process is considered. A new two-staged algorithm for their estimation is developed. In the first stage initial estimates of the intensity parameters are constructed using the method of moments and maximum likelihood method. In the second stage these estimates are used in the cumulative sums change-point detection algorithm to determine time intervals, during which the process intensity is constant; after that estimates of all parameters are constructed. The developed algorithm is compared with a well-known algorithm for finding maximum likelihood estimates (the EM algorithm). The proposed method is non-iterative, ensuring relatively low computational complexity.

Keywords: Markov modulated Poisson process; MMPP; cumulative sums; method of moments; maximum likelihood method.

Acknowledgements. This work was partially carried out with the financial support of the Russian Foundation for Basic Research (grant No. 24-11-00191 in the name of S. E. Vorobeychikov).

Introduction

Stochastic processes with changing parameters, in particular, processes with change points, are widely used for approximating real non-linear processes in different applications (finance, medicine, psychology, geology, literature and so on) and even in our daily lives. In the context of statistics a change point is a place or time point such that the observations follow one distribution up to that point and follow another distribution after that point. Multiple change points problems can also be defined similarly [1]. Processes with a single change point are usually used to simulate a withdraw from stable conditions or an equipment error, whereas processes with multiple change points describe real processes with several states, where switches between states occur in unknown instants. The mentioned above book is devoted to parametric change-point detection with applications to genetics, finance and medicine and contains a number of classical and recent results in this field for some commonly used distributions.

In queue theory processes with multiple change points are applied to describe queue systems with different regimes of customers arrivals. A process with two states corresponding to «usual» and «peak» time can be considered as the simplest model. These two states are characterised of different rates of arrivals: at the first case events occur in general much rarely then in the second case. When the switching times are controlled by a Markov chain and the flow of events have Poisson distribution, the model is named as Markov modulated Poisson process (MMPP). The time intervals between events can be considered as random variables obeyed to a mixture of two or more distributions.

Various modifications of the expectation maximisation (EM) algorithm based on maximum likelihood estimates are usually used to estimate mixtures of distributions. Each iteration of the algorithm consists of two steps. In the expectation step the expected value of the likelihood function is calculated using the current parameter estimate. In the maximisation step the parameters that maximise the expected maximum likelihood function are calculated. For the first time an algorithm with this name was proposed in paper [2], although iterative maximum likelihood estimates had been proposed earlier: for example, in article [3] they were used for the analysis of Markov chains. Later a large number of publications appeared in which the proposed approach was applied and modified for different classes of random processes, including processes characterised by a mixture of exponential distributions.

Markovian arrival processes have been introduced in 1970s and since then they and, in particular, MMPP have extensively been used as models for input streams to queueing systems. In paper [4] MMPP is applied to web traffic modelling. Article [5] uses MMPP to describe human inter-contact times. In work [6] a modified MMPP is applied for modelling disease dynamics based on medical claims data.

Paper [7] can be mentioned as one of the first devoted to EM estimates for MMPP. It presents a detailed description of the algorithm and some numerical examples. The author stresses that the algorithm demands some complex calculations and it should be taken into account to determine the effectiveness; nevertheless, it demands much less iterations than others. Article [8] contains a specification of the EM algorithm in order to

fit an empirical counting process observed at discrete times to a Markovian arrival process. An exact numerical procedure to compute the EM steps is introduced and some examples are given. Work [9] examines maximum likelihood estimation of the second order Markovian arrival processes parameterisation of the two-state Markovian arrival processes. The problems of the starting solution and of the carrying out the numerical evaluation of the likelihood function are stated and discussed. For the starting solution the method of moments can be applied.

A review of various modifications of the EM algorithm is given in paper [10]. The positive aspects of these algorithms are due to the fact that they are based on the maximum likelihood method, which guarantees the consistency and asymptotic efficiency of estimates. However, a number of works note the disadvantages of EM algorithms. Namely, the convergence of the EM algorithm is proved under the mandatory condition of the boundedness of the log-likelihood function. Also, the presence of a large number of local maxima of the log-likelihood function leads to a large instability with respect to the initial approximation and the initial data [11]. In addition, the iterative nature of the algorithms leads to a large number of calculations, the volume of which increases with a significant number of parameters [12]. In this regard alternative approaches to estimating the parameters of mixture distributions are also being developed. In work [12] a mixture of exponential distributions is used to approximate distributions with heavy tails, which are characteristic of file lengths, call retention times, the lengths of scenes in moving picture experts group video streams and the intervals between connection requests in Internet traffic. The authors demonstrate the feasibility of approximating a distribution with heavy tails by a hyperexponential distribution and propose a recursive fitting algorithm. Firstly, the minimum intensity, which provides the highest values of the random variable and thus contributes the most to the tail of the distribution, is estimated. Secondly, this distribution is subtracted from the mixture distribution, and the intensity is estimated. Article [13] develops this approach for approximating the Pareto distribution, when estimating the average duration of the cumulative sums procedure for which an explicit formula is obtained as a result. In paper [14] the authors use a hyperexponential distribution to describe sensor networks. For estimating the parameters two original algorithms that minimise the objective function are proposed. The first algorithm consists of two stages. In the first stage the intensity parameters are estimated, in the second stage the probabilities of occurrence of observations with given intensities. Probabilities are expressed through intensities, which allows to reduce the number of estimated parameters by half at each stage. The second algorithm estimates all parameters simultaneously.

In paper [15] we consider the problem of estimating the parameters of a hyperexponential distribution with two values of the intensity. Explicit formulas for moment estimates were obtained. Also, the combined parameter estimates that use the first two sample moments in the likelihood function are proposed. In this paper we develop this approach. The properties of the combined estimates are analysed and used to evaluate MMPP parameters. In work [16] we developed a cumulative sum algorithm for Poisson process change-point detection, when parameters of the process are known. The estimates of the parameters instead of their values are used.

Problem statement

MMPP is a doubly stochastic Poisson process, the rate of which varies according to a Markov chain. We consider a process with two rates determined by the states of a non-observable controlling Markov chain. The sojourn time in state $\{j\}$, $j \in \{1, 2\}$, is determined by exponentially distributed random variables $\{x_i\}$, $i = 1, 2, \dots$, with parameters μ_j , where x_i for odd i and for even i are distributed with different parameters.

Let us denote $x_1 + \dots + x_m$ as T_m and $T_0 = 0$. At interval $[T_{i-1}, T_i)$ we can observe a Poisson process with the rate λ_j , where j is the state of the controlling Markov chain at the interval. The instants T_i can be considered as the change points of the observed Poisson process, where the rate of the process changes. We denote the arrival times of the observed process as t_k , $k = 1, 2, \dots$, $t_0 = 0$.

The problem is to estimate parameters $\{\lambda_1, \lambda_2, \mu_1, \mu_2\}$ from observations t_k .

Parameter estimation

We consider the process $\{\tau_i\}_{i \geq 1}$, where $\tau_i = t_i - t_{i-1}$ is the length of the i interval between arriving events in the observed flow. The values $\{\tau_i\}$ for MMPP are independent and exponentially distributed with one of two possible parameters. It allows us to use the hyperexponential distribution as a model for the observed data. The hyperexponential distribution has the density function

$$f(\tau) = p\lambda_1 e^{-\lambda_1 \tau} + (1-p)\lambda_2 e^{-\lambda_2 \tau}, \quad p \in (0, 1), \quad (1)$$

which is a mixture of two exponential distributions.

To estimate parameters λ_1, λ_2 and p of the hyperexponential distribution we can apply either the method of moments (using three sample moments) or the maximum likelihood method. In the latter case maximisation of the likelihood relation demands the use of numerical methods with the prescribed accuracy.

We propose combined estimates on the basis of both mentioned above methods. Let us have a sample $\{\tau_1, \dots, \tau_N\}$ obeyed distribution (1). Primarily, we calculate the following two first moments:

$$m_1 = \frac{1}{N} \sum_{k=1}^N \tau_k, \quad m_2 = \frac{1}{N} \sum_{k=1}^N \tau_k^2.$$

Then, equating them to the corresponding theoretical moments, we have the system of equations

$$\begin{cases} m_1 = \frac{p}{\lambda_1} + \frac{1-p}{\lambda_2}, \\ m_2 = \frac{2p}{\lambda_1^2} + \frac{2(1-p)}{\lambda_2^2}. \end{cases}$$

Let us introduce the following notations:

$$x = \frac{p}{\lambda_1}, \quad y = \frac{1-p}{\lambda_2}. \quad (2)$$

We obtain the system of equations

$$\begin{cases} m_1 = x + y, \\ m_2 = \frac{2x^2}{p} + \frac{2y^2}{1-p}. \end{cases}$$

Expressing from the first equation $y = m_1 - x$ and substituting it in the second equation, we come to the quadratic equation

$$x^2 - 2m_1px + p\left(m_1^2 - \frac{1-p}{2}m_2\right) = 0.$$

If the inequality $m_2 - 2m_1^2 \geq 0$ holds true, the equation has two solutions:

$$x_{1,2} = pm_1 \pm \sqrt{p(1-p)} \sqrt{\frac{m_2}{2} - m_1^2}.$$

Choosing the largest root, we have

$$\begin{aligned} x &= pm_1 + \sqrt{p(1-p)} \sqrt{\frac{m_2}{2} - m_1^2}, \\ y &= (1-p)m_1 - \sqrt{p(1-p)} \sqrt{\frac{m_2}{2} - m_1^2}. \end{aligned}$$

Then using the notations (2), values λ_1, λ_2 are defined as functions of the parameter p :

$$\begin{aligned} \hat{\lambda}_1 &= \frac{1}{m_1 + \sqrt{\frac{1-p}{p}} \sqrt{\frac{m_2}{2} - m_1^2}}, \\ \hat{\lambda}_2 &= \frac{1}{m_1 - \sqrt{\frac{p}{1-p}} \sqrt{\frac{m_2}{2} - m_1^2}}. \end{aligned} \quad (3)$$

It should be noted that $\hat{\lambda}_1 < \hat{\lambda}_2$. If we take the smallest root for x , it obtains the same estimators for λ_1, λ_2 , except of $\hat{\lambda}_2 < \hat{\lambda}_1$.

As both rates should be positive, parameter p should obey the following inequality:

$$m_1 - \sqrt{\frac{p}{1-p}} \sqrt{\frac{m_2}{2} - m_1^2} > 0.$$

It means that

$$p < \frac{2m_1^2}{m_2}.$$

Then we define the log-likelihood function

$$L(p) = \sum_{n=1}^N \log(p\lambda_1 e^{-\lambda_1 \tau_n} + (1-p)\lambda_2 e^{-\lambda_2 \tau_n}),$$

where parameters $\{\lambda_1, \lambda_2\}$ are calculated according to equations (3). Maximising the function $L(p)$ on p , $0 < p < \frac{2m_1^2}{m_2}$, and using relations of the equations (3), we obtain the estimates of parameters $\{\lambda_1, \lambda_2\}$.

In paper [16] we developed a cumulative sum algorithm for Poisson process change point detection, when parameters $\{\lambda_1, \lambda_2\}$ are known. To construct algorithm we use statistics and cumulative sums from work [16]. Statistics z_i is defined as the logarithm of the likelihood ratio for exponential distribution

$$z_i = \ln \left(\frac{\lambda_2 \exp(-\lambda_2 \tau_i)}{\lambda_1 \exp(-\lambda_1 \tau_i)} \right) = \ln \frac{\lambda_2}{\lambda_1} - (\lambda_2 - \lambda_1) \tau_i. \quad (4)$$

Let $\lambda(t)$ be the rate of the arrival process at moment t . We introduce the following two hypothesis:

$$H_j = \{\lambda(t) = \lambda_j, j = 1, 2\}.$$

We suppose that $\lambda_1 < \lambda_2$ and denote $\delta = \frac{\lambda_2}{\lambda_1}$, where $\delta > 1$. Statistics z_i in distribution (4) have the form

$$z_i = \ln \delta - \lambda_1 (\delta - 1) \tau_i.$$

Under hypothesis H_j , these statistics have the following properties:

$$E[z_i | H_1] = \ln \delta - (\delta - 1) < 0,$$

$$E[z_i | H_2] = \ln \delta - \left(\frac{\delta - 1}{\delta} \right) > 0.$$

Note that the mean of statistics z_i changes the sign with the change the rate of the process.

In our algorithm we use estimators (3) obtained above instead of the parameter values. We introduce positive values $h^{(1)}, h^{(2)}$ as the algorithms thresholds and construct cumulative sum S_i , which is recalculated at the instants of arrivals:

$$S_0 = 0, S_i = \begin{cases} \max\{0, S_{i-1} + c_{i-1} z_i\}, & \text{if } S_{i-1} + c_{i-1} z_i < h_i, i \geq 1, \\ 0, & \text{if } S_{i-1} + c_{i-1} z_i \geq h_i, i \geq 1, \end{cases}$$

$$z_i = \ln \frac{\hat{\lambda}_2}{\hat{\lambda}_1} - (\hat{\lambda}_2 - \hat{\lambda}_1) \tau_i,$$

$$c_0 = 1, c_i = \begin{cases} c_{i-1}, & \text{if } S_{i-1} + c_{i-1} z_i < h_i, i \geq 1, \\ -c_{i-1}, & \text{if } S_{i-1} + c_{i-1} z_i \geq h_i, i \geq 1, \end{cases} \quad (5)$$

$$h_i = \begin{cases} h^{(1)}, & \text{if } c_{i-1} = 1, \\ h^{(2)}, & \text{if } c_{i-1} = -1. \end{cases}$$

If $c_i = 1$, the sum detects the increase of the rate, i. e. the change of the hypothesis from H_1 to H_2 , when it reaches threshold $h^{(1)}$. If $c_i = -1$, the sum detects the decrease of the rate, i. e. the change of the hypothesis from H_2 to H_1 , when it reaches threshold $h^{(2)}$.

Let sequence $\{\sigma_m\}_{m \geq 0}$ be the sequence of instants, when the cumulative sum reaches the threshold h_i , i. e.

$$\sigma_0 = 0, \sigma_m = \min\{t_i > \sigma_{m-1} : S_i \geq h\}, m \geq 1. \quad (6)$$

Consider sequence $\{n_m\}_{m \geq 0}$ associated with sequence $\{\sigma_m\}_{m \geq 0}$ as follows:

$$n_0 = 0, n_m = \max\{t_i \leq \sigma_m : S_i > 0, S_{i-1} = 0\}, m \geq 1. \quad (7)$$

Thus, instant n_m is the first instant, when the cumulative sum becomes positive to reach then the threshold.

The algorithm for change-point detection is described as follows. Calculate the cumulative sum given by equation (5). Then construct sequences $\{\sigma_m\}$, $\{n_m\}$ defined by equations (6) and (7). If for the last instant n_m one has $m = 2l + 1$, one sets $n_{2l+2} = t_N$. If for the last instant n_m one has $m = 2l$, one sets $n_{2l+1} = t_N$. Here the odd instants n_{2l+1} are the estimators of the instants, when the rate changes from λ_1 to λ_2 , and the even instants n_{2l+2} are the estimators of the instants when the rate changes from λ_2 to λ_1 . So, we consider the Markov chain to be in the first state at intervals $[n_{2l}, n_{2l+1}]$ and in the second state at intervals $[n_{2l+1}, n_{2l+2}]$.

According to sequential change-point detection algorithms two types of errors are considered: the false alarm and the skipping of the change. A false alarm occurs, when one of the cumulative sums reaches the corresponding threshold in the case of the constant intensity of the arrival process. A skipping of the change occurs, when the change of the parameter occurs but the corresponding cumulative sum did not reach its threshold. Usual characteristics of any sequential algorithm are the mean delay and the mean length of the interval between false alarms, as well as the probabilities of a false alarm and of a skipping the change.

When implementing the algorithm, it is possible to encounter false alarm situations. We shall record all the exceeding the thresholds by either the first or the second cumulative sum. If the same sum reaches threshold several times in a row, we only record the first occurrence.

Now we are ready to estimate parameters $\{\mu_1, \mu_2\}$ of the controlling Markov chain and parameters $\{\lambda_1, \lambda_2\}$.

Let N_1 be the number of intervals $[q_{2l}, q_{2l+1}]$ and T_1 be the total length of these intervals. Also, let N_2 be the number of intervals $[q_{2k+1}, q_{2k+2}]$ and T_2 be the total length of these intervals. The estimates for $\{\mu_1, \mu_2\}$ have the following form:

$$\hat{\mu}_j = \frac{N_j}{T_j}, j = 1, 2. \quad (8)$$

Let M_1 be the number of events at intervals $[q_{2l}, q_{2l+1}]$. Also, let M_2 be the number of events at intervals $[q_{2k+1}, q_{2k+2}]$. Thus, we can estimate rates $\{\lambda_1, \lambda_2\}$ as reciprocal values to the mean length of the interval between neighbouring events, when the controlling chain is in the corresponding state:

$$\hat{\lambda}_j = \frac{M_j}{T_j}, j = 1, 2. \quad (9)$$

Simulation results and their discussion

To analyse the algorithm simulations were conducted. The results are represented in table 1. Here $\{\lambda_1, \lambda_2, \mu_1, \mu_2\}$ are the process parameters, $\{\hat{\lambda}_1, \hat{\lambda}_2, \hat{\mu}_1, \hat{\mu}_2\}$ are the estimates (8) and (9), h_1, h_2 are the thresholds of the change-point detection algorithm. The parameters in first two rows were chosen according to paper [7], where they were used to implement EM algorithm.

Table 2 represents the simulation results from article [7]. Here I is the number of iterations, other notations are the same as in table 1. Columns for estimates $\{\hat{\lambda}_1, \hat{\lambda}_2, \hat{\mu}_1, \hat{\mu}_2\}$ have two rows, the first one contains initial values of parameters $\{\lambda_1, \lambda_2, \mu_1, \mu_2\}$, which are necessary for EM algorithm. It should be mentioned that they are not calculated in the frame of the algorithm.

Table 1

Simulation results

λ_1	λ_2	$\hat{\lambda}_1$	$\hat{\lambda}_2$	μ_1	μ_2	$\hat{\mu}_1$	$\hat{\mu}_2$	h_1	h_2
50	100	46.72	105.96	2.0	5.0	2.160	4.970	1.0	4.7
10	100	9.95	106.39	1.0	10.0	1.010	10.610	2.5	4.1
50	100	49.46	102.19	1.0	2.0	1.10	2.070	1.3	4.4
1	2	0.98	2.16	0.04	0.1	0.038	0.090	1.0	5.3
1	10	1.02	10.63	0.10	1.0	0.097	1.050	2.2	4.1
1	3	0.96	3.13	0.04	0.1	0.045	0.096	2.5	2.9

Table 2

Simulation results for EM algorithm

λ_1	λ_2	$\hat{\lambda}_1$	$\hat{\lambda}_2$	μ_1	μ_2	$\hat{\mu}_1$	$\hat{\mu}_2$	I
50	100	53.9	128	2	5	1.90	10.12	26
50	100	50.8	105	2	5	2.94	5.26	26

The results demonstrate a good quality of the proposed approach. The error of the estimates is similar to EM estimates but the procedure is much simpler because it is not iterative.

Conclusions

A new algorithm for estimation of the parameters of a second order Markov modulated Poisson process is proposed. Simulation results demonstrate the algorithm quality, which is similar to widely used EM algorithm. The proposed approach does not demand the initial estimates for all parameters and it is not iterative. Hence, the computational complexity of the algorithm is much less than the computational complexity of the EM method. We are going to develop our approach for arbitrary order MMPP.

References

1. Chen J, Gupta AK. *Parametric statistical change point analysis with applications to genetics, medicine, and finance*. 2nd edition. Boston: Birkhäuser; 2012. XIII, 273 p.
2. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B, Methodological*. 1977;3(1):1–38.
3. Baum LE, Petrie T, Soules G, Weiss N. A maximization technique occurring in the statistical analysis of probabilistic function of Markov chains. *Annals of Mathematical Statistics*. 1970;41(1):164–171.
4. Scott SL, Smyth P. The Markov modulated Poisson process and Markov poisson cascade with applications to web traffic modeling. In: Lindley V, Bernardo JM, Bayarri MJ, Berger JO, Dawid AP, Heckerman D, editors. *Bayesian statistics 7. Proceedings of the Seventh Valencia international meeting; 2002 June 2–6; Oxford, UK*. Oxford: Oxford University Press; 2003. p. 671–680.
5. Frohn S, Gübner S, Lindemann Ch. An accurate and analytically tractable model for human inter-contact times. In: Pazzi RW, editor. *Proceedings of the International conference on modeling, analysis and simulation of wireless and mobile systems; 2010 October 17–21; Bodrum, Turkey*. New York: Association for Computing Machinery; 2010. p. 275–282. DOI: 10.1145/1868521.1868565.
6. Mews S, Surmann B, Hasemann L, Elkenkamp S. Markov modulated marked Poisson processes for modelling disease dynamics based on medical claims data. *Statistics in Medicine*. 2023;42(21):3804–3815. DOI: 10.1002/sim.9832.
7. Ryden T. An EM algorithm for estimation in Markov modulated Poisson processes. *Computational Statistics & Data Analysis*. 1996;21(4):431–447.
8. Breuer L, Kume A. An EM algorithm for Markovian arrival processes observed at discrete times. In: Müller-Clostermann B, Echtle K, Rathgeb EP, editors. *Proceedings of the conference on measurement, modelling, and evaluation of computing systems and dependability and fault tolerance; 2010 March 15–17; Essen, Germany*. Berlin: Springer; 2010. p. 242–258.
9. Carrizosa E, Ramirez-Cobo P. Maximum likelihood estimation in the two-state Markovian arrival process. arXiv:14013105 [Preprint]. 2014 [cited 2025 December 10]: [20 p.]. Available from: <https://arxiv.org/pdf/1401.3105>.
10. Korolev VYu. *EM algorithm, its modifications and their application to the problem of separation of mixtures of probability distributions. Theoretical review*. Moscow: Institute of Informatics Problems of the Russian Academy of Sciences; 2007. 102 p. Russian.
11. Korolev VYu, Nazarov AL. Separation mixtures of probability distributions with the grid method of moments and the grid maximum likelihood method. *Automation and Remote Control*. 2010;71(3):455–472.

12. Feldmann A, Whitt W. Fitting mixtures of exponentials to long-tail distributions to analyze network performance models. *Performance Evaluation*. 1998;31(3–4):245–279.
13. Petcharat K, Areepong Yu, Sukparungsee S, Mititelu G. Fitting Pareto distribution with hyperexponential to evaluate the ARL for CUSUM chart. *International Journal of Pure and Applied Mathematics*. 2012;77(2):233–244.
14. Singh LN, Dattatreya GR. Estimation of the hyperexponential density with applications in sensor networks. *International Journal of Distributed Sensor Networks*. 2002;3(3):311–330.
15. Burkatovskaya YuB, Vorobeychikov SE. Parameter estimation for hyperexponential distribution. *Tomsk State University Journal of Control and Computer Science*. 2025;70:52–61. Russian. DOI: 10.17223/19988605/70/5.
16. Burkatovskaya Y, Kabanova T, Vorobeychikov S. CUSUM algorithms for parameter estimation in queueing systems with jump intensity of the arrival process. In: Dudin A, Nazarov A, Yakupov R, editors. *Information technologies and mathematical modeling – queueing theory and applications. Proceedings of the 14th International scientific conference; 2015 November 18–22; Anzhero-Sudzhensk, Russia*. Cham: Springer; 2015. p. 275–288.

Received 24.12.2025 / revised 19.01.2026 / accepted 29.01.2026.

ЛОКАЛЬНАЯ ИНФОРМАЦИОННАЯ ГЕОМЕТРИЯ ДЛЯ СТАТИСТИЧЕСКОГО АНАЛИЗА ДВОИЧНЫХ ЦЕПЕЙ МАРКОВА ВЫСОКОГО ПОРЯДКА

В. А. ВОЛОШКО^{1), 2)}

¹⁾Белорусский государственный университет, пр. Независимости, 4, 220030, г. Минск, Беларусь

²⁾Научно-исследовательский институт прикладных проблем математики и информатики БГУ,
пр. Независимости, 4, 220030, г. Минск, Беларусь

Аннотация. На 14-й Международной научной конференции «Компьютерный анализ данных и моделирование: стохастика и анализ данных» автором были представлены некоторые новые результаты относительно асимптотических свойств статистик от чисто случайной (равномерно распределенной) двоичной последовательности растущей длины. Эти результаты получены с применением методов информационной геометрии к многообразиям марковских распределений вероятностей на множестве бесконечных двоичных последовательностей. В данной статье более детально описываются информационно-геометрическая теория и техника доказательств. На нескольких примерах показывается, как нахождение вероятностно-статистических свойств сводится к геометрическим и комбинаторным вычислениям.

Ключевые слова: информационная геометрия; касательное пространство; асимптотическое распределение; двоичная последовательность; марковское многообразие.

LOCAL INFORMATION GEOMETRY FOR STATISTICAL ANALYSIS OF HIGH-ORDER BINARY MARKOV CHAINS

V. A. VOLOSHKO^{a, b}

^aBelarusian State University, 4 Niezaliezhnasci Avenue, Minsk 220030, Belarus

^bResearch Institute for Applied Problems of Mathematics and Informatics,
Belarusian State University, 4 Niezaliezhnasci Avenue, Minsk 220030, Belarus

Abstract. At the 14th International conference «Computer data analysis and modelling: stochastics and data analysis» the author presented some new results on asymptotic properties of statistics derived from purely random (uniformly distributed) binary sequence of increasing length. These results are obtained by methods of information geometry applied to manifolds of Markov probability distributions on the set of infinite binary sequences. In this paper the underlying information-geometric theory and the technics of proofs are describe in more detail. By few examples, it is shown how finding probabilistic and statistical properties comes down to geometric and combinatorial computations.

Keywords: information geometry; tangent space; asymptotic distribution; binary sequence; Markov manifold.

Образец цитирования:

Волошко ВА. Локальная информационная геометрия для статистического анализа двоичных цепей Маркова высокого порядка. *Журнал Белорусского государственного университета. Математика. Информатика.* 2026;1:61–74 (на англ.). EDN: OGPXRN

For citation:

Voloshko VA. Local information geometry for statistical analysis of high-order binary Markov chains. *Journal of the Belarusian State University. Mathematics and Informatics.* 2026;1:61–74. EDN: OGPXRN

Автор:

Валерий Анатольевич Волошко – кандидат физико-математических наук; доцент кафедры математического моделирования и анализа данных факультета прикладной математики и информатики¹⁾, заведующий сектором компьютерного анализа данных²⁾.

Author:

Valeriy A. Voloshko, PhD (physics and mathematics); associate professor at the department of mathematical modelling and data analysis, faculty of applied mathematics and computer science^a, and head of the sector of computer data analysis^b.
valoshka@bsu.by
<https://orcid.org/0000-0002-9693-0688>



Markov manifolds and tangent spaces

Let $x_1, \dots, x_n \in Y := \{0, 1\}$ be purely random (uniformly distributed) binary sequence of length $n \in \mathbb{N} : P\{x_1^n = q\} \equiv 2^{-n}$, $q \in Y^n$, $x_1^n := (x_i)_{i=1}^n$. For convenience, we consider index i of x_i as discrete time and use corresponding «temporal» terminology. The new results on asymptotic properties of statistics derived from x_1^n as $n \rightarrow +\infty$ presented in article [1] are obtained by methods of information geometry [2] applied to manifolds of Markov probability distributions (Markov manifolds for brevity) on the set $Y^{\mathbb{Z}} := \{x = (x_i)_{i \in \mathbb{Z}} : x_i \in Y\}$ of two-sided infinite binary sequences.

Introduce some notations: $a | b$ means « a divides b »; $\ell\{\cdot\}$ is the probability distribution of a random variable; $\text{cov}\{\cdot, \cdot\}$, $\text{cor}\{\cdot, \cdot\}$ are the covariance and the correlation coefficient for a pair of random variables, respectively; « $\lim_{\text{cov}}$ » and « $\lim_{\text{cor}}$ » briefly denote their limits as $n \rightarrow +\infty$; $\mathbb{Z}_m = \{0, \dots, m-1\}$ is the ring of integers modulo $m \in \mathbb{N}$ standardly associated with a non-negative integer interval; $a \stackrel{m}{=} b$ and $a \stackrel{m}{+} b$ are equality and summation modulo m , respectively; $\text{span}\{\cdot\}$ is a linear span of subset of linear space; $\text{lcd}(\cdot, \cdot)$, $\text{gcd}(\cdot, \cdot)$ are lowest common dominator and greatest common divisor, respectively; $[a : b] = \{a, a+1, \dots, b\}$, $a, b \in \mathbb{Z}$, $a \leq b$.

We consider two nested Markov manifolds.

The first Markov manifold $\mathbb{M}$ consists of distributions of stationary Markov chains of all orders $s \in \mathbb{N}_0 = \{0, 1, \dots\}$ determined by stationarity condition and Markov equation:

$$P\{x_i = 0 | x_{i-j} = q_j, j \in \mathbb{N}\} = \frac{1 + \omega(q)}{2}, q = (q_j)_{j \in \mathbb{N}} \in Y^{\mathbb{N}}, i \in \mathbb{Z}, \quad (1)$$

where the function $\omega : Y^{\mathbb{N}} \rightarrow \mathbb{R}$ depends only on finite number of components $\{q_j\}$ and $\|\omega\|_{\infty} = \max_{q \in Y^{\mathbb{N}}} |\omega(q)| < 1$.

We call order of ω and denote by $\text{ord}(\omega)$ the minimal value $s \in \mathbb{N}_0$ such that $\omega(q)$ does not depend on $\{q_j : j > s\}$. Manifold $\mathbb{M}$ is in a bijection with the set of functions ω , so we identify them as

$$\mathbb{M} = \{\omega : \text{ord}(\omega) < +\infty, \|\omega\|_{\infty} < 1\}.$$

The uniform distribution $u \in \mathbb{M}$ corresponds to zero function $\omega(q) \equiv 0$, $q \in Y^{\mathbb{N}}$. If we take off the norm limitation $\|\omega\|_{\infty} < 1$, we get the model of tangent space [2] (linearised neighbourhood) of uniform distribution u within the Markov manifold $\mathbb{M}$:

$$\mathbb{T} = \{\omega : \text{ord}(\omega) < +\infty\}.$$

The fundamental object of information geometry defined on this tangent space is the Fisher – Riemann dot product [2]. It has the following form in our notation:

$$\langle \omega, \omega' \rangle = 2^{-s} \sum_{q \in Y^s} \omega(q \| 0^{\infty}) \omega'(q \| 0^{\infty}), \omega, \omega' \in \mathbb{T}, s = \max\{\text{ord}(\omega), \text{ord}(\omega')\},$$

where $\|$ is concatenation; $0^{\infty} \in Y^{\mathbb{N}}$ is one-sided zero binary sequence. We call elements $\omega \in \mathbb{T}$ tangent vectors that is the standard name for elements of tangent spaces.

Under local information geometry (locality refers to neighbourhood of u) we understand here the structure of tangent space $\mathbb{T}$ with Fisher – Riemann dot product on it, the objects defined on this structure and the meaningful quantities that can be computed based on it.

An orthonormal base of characters [3] $\chi = \{\chi^{(Q)}\}_{Q \in Y^+}$ is defined on $\mathbb{T}$:

$$\chi^{(Q)} \in \mathbb{T}, \chi^{(Q)}(q) := e\left(\bigoplus_{j \in \mathbb{N}} Q_j q_j\right), Q \in Y^+, q \in Y^{\mathbb{N}}, \quad (2)$$

where $Y^+ \subset Y^{\mathbb{N}}$ is the subset of one-sided binary sequences of the finite Hamming weight $\|Q\| := \sum_{j \in \mathbb{N}} Q_j < +\infty$

(with finite number of ones); $e(z) := (-1)^z$ is a discrete exponent. Infinite XOR in formula (2) is correct because it contains only a finite number of ones due to $Q \in Y^+$. Note that Y^+ is a countable subset of a continuous set $Y^{\mathbb{N}}$. Many important tangent subspaces $T \subset \mathbb{T}$ are linear spans of subsets of the characters base (2): we call χ -canonic such subspaces T . The notion of order is naturally transferred from $\mathbb{T}$ to Y^+ by formula (2):

$$\text{ord}(Q) := \text{ord}(\chi^{(Q)}), Q \in Y^+, Y^+[s] := \{Q \in Y^+ : \text{ord}(Q) = s\}, s \in \mathbb{N}_0. \quad (3)$$

Zero sequence has zero order: $\text{ord}(0^\infty) = 0$. For non-zero sequences $Q \in Y^+$, $\|Q\| > 0$, the order is the position of the rightmost unit: $Q_{\text{ord}(Q)} = 1$, $Q_i \equiv 0$, $i > \text{ord}(Q)$.

Submanifolds $M(s) \subset M$ and subspaces $T(s) \subset T$ are defined as follows:

$$M(s) := \{\omega \in M : \text{ord}(\omega) \leq s\}, \quad T(s) := \{\omega \in T : \text{ord}(\omega) \leq s\}, \quad s \in \mathbb{N}_0.$$

Obviously $M(s) \subset M(s')$, $T(s) \subset T(s')$ for any $0 \leq s < s'$. Space $T(s)$ is the tangent space of u within the manifold $M(s)$. Denote by $\partial T(s)$, $s > 0$, an orthogonal complement of $T(s-1)$ within $T(s)$ with regard to the Fisher – Riemann dot product. For $s = 0$ set $\partial T(0) = T(0)$. Spaces $T(s)$, $\partial T(s)$ are χ -canonic and have the following dimensions:

$$\dim(T(s)) = 2^s = \sum_{s'=0}^s \gamma[s'], \quad \dim(\partial T(s)) = \gamma[s] := 2^{\max\{0, s-1\}}, \quad s \in \mathbb{N}_0. \quad (4)$$

Base of characters for $\partial T(s)$ is $\{\chi^{(Q)} : Q \in Y^+[s]\}$, $\gamma[s] = |Y^+[s]|$ (see formulas (3) and (4)).

Let us call random sequence $x \in Y^{\mathbb{Z}}$: Δ -stationary if its probability distribution is invariant under shift by $\Delta \in \mathbb{N}$, i. e. $\ell(x_i)_i = \ell(x_{i+\Delta})_i$; semi-stationary if it is Δ -stationary for some $\Delta \in \mathbb{N}$. «Usual» stationarity is 1-stationarity in these terms.

The second Markov manifold that we consider is $M^\pi = \bigcup_{\Delta \in \mathbb{N}} M^\pi\{\Delta\}$, where the submanifold $M^\pi\{\Delta\} \subset M^\pi$ consists of distributions of Δ -periodic Δ -stationary Markov chains determined by Δ -stationarity condition and Δ -periodic generalisation of Markov equation (1):

$$P\{x_i = 0 | x_{i-j} = q_j, j \in \mathbb{N}\} = \frac{1 + \omega_i(q)}{2}, \quad q \in Y^{\mathbb{N}}, \quad i \in \mathbb{Z}, \quad (5)$$

$\omega_i : \mathbb{Z} \rightarrow M$ is Δ -periodic with regard to $i \in \mathbb{Z}$ M -valued sequence: $\omega_i \equiv \omega_{i+\Delta}$. We identify the set of such sequences with $M^\pi\{\Delta\}$, and with M^π for all $\Delta \in \mathbb{N}$. Denote by T^π the tangent space of u within M^π . Obviously $M \subset M^\pi$ and $T \subset T^\pi$. The model of T^π is the set of periodic T -valued sequences $\omega_i : \mathbb{Z} \rightarrow T$, where the subspace $T^\pi\{\Delta\} \subset T^\pi$ of Δ -periodic ones forms the tangent space of u within $M^\pi\{\Delta\}$ (we use braces $M^\pi\{\cdot\}$, $T^\pi\{\cdot\}$ to avoid confusion with the notations $M(\cdot)$, $T(\cdot)$). Spaces $\{T^\pi\{\Delta\}\}$ are nested by periods divisibility: $T^\pi\{\Delta\} \subset T^\pi\{\Delta'\}$ for any $\Delta | \Delta'$ (similarly for manifolds $\{M^\pi\{\Delta\}\}$). In particular, $T^\pi\{1\} \subset T^\pi\{\Delta\}$ for any $\Delta \in \mathbb{N}$, where $T^\pi\{1\}$ is the subspace of constant sequences $\omega_i \equiv \text{const}_i$. This subspace implements the embedding $T \rightarrow T^\pi$, i. e. $T^\pi\{1\}$ is T embedded into T^π . Denote by $\text{per}(\omega)$ the minimal period $\min\{\Delta \in \mathbb{N} : \omega \in T^\pi\{\Delta\}\}$ of $\omega \in T^\pi$. Fisher – Riemann dot product on T^π has the form

$$\langle \omega, \omega' \rangle = \Delta^{-1} \sum_{i=1}^{\Delta} \langle \omega_i, \omega'_i \rangle, \quad \omega, \omega' \in T^\pi, \quad \Delta = \text{lcd}(\text{per}(\omega), \text{per}(\omega')),$$

where $\langle \omega_i, \omega'_i \rangle$ is the Fisher – Riemann dot product on T .

The space T^π is representable as a tensor product

$$T^\pi = T \otimes \Pi, \quad (6)$$

where $\Pi = \bigcup_{\Delta \in \mathbb{N}} \Pi\{\Delta\}$ is a space of two-sided periodic real-valued sequences $a_i : \mathbb{Z} \rightarrow \mathbb{R}$, $\Pi\{\Delta\}$ is its subspace of Δ -periodic sequences $a_i \equiv a_{i+\Delta}$. By tensor product $U \otimes V$ of two spaces, we mean the linear span of tensor product $\{u_i\} \otimes \{v_j\} = \{u_i \otimes v_j\}$ (pairwise) of their bases $\{u_i\} \subset U$, $\{v_j\} \subset V$. Tensor product $\tilde{\omega} = \omega \otimes a \in T^\pi$ of two single elements $\omega \in T$ and $a \in \Pi$ is

$$\tilde{\omega} = (\tilde{\omega}_i)_{i \in \mathbb{Z}}, \quad \tilde{\omega}_i = a_i \omega. \quad (7)$$

In these terms $T^\pi\{\Delta\} = T \otimes \Pi\{\Delta\}$, $\Delta \in \mathbb{N}$. On the space Π a dot product similar to representation (6) is defined as

$$\langle a, a' \rangle = \Delta^{-1} \sum_{i=1}^{\Delta} \langle a_i, a'_i \rangle, \quad a, a' \in \Pi, \quad \Delta = \text{lcd}(\text{per}(a), \text{per}(a')), \quad (8)$$

and under tensor product (7) dot products on $\mathbb{T}$ and on Π are multiplied:

$$\langle \tilde{\omega}, \tilde{\omega}' \rangle = \langle \omega, \omega' \rangle \langle a, a' \rangle, \quad \omega, \omega' \in \mathbb{T}, \quad a, a' \in \Pi, \quad \tilde{\omega} = \omega \otimes a, \quad \tilde{\omega}' = \omega' \otimes a'. \quad (9)$$

Denote canonic orthonormal base $\psi^{(\Delta)} = \left\{ \psi^{(t, \Delta)} \right\}_{t \in \mathbb{Z}_\Delta}$ of $\Pi \{ \Delta \}$:

$$\psi^{(t, \Delta)} \in \Pi \{ \Delta \}, \quad \psi_i^{(t, \Delta)} = \sqrt{\Delta} \cdot 1 \left\{ i \equiv t \right\}, \quad \Delta \in \mathbb{N}, \quad t \in \mathbb{Z}_\Delta. \quad (10)$$

Note that the combined set $\psi = \bigcup_{\Delta \in \mathbb{N}} \psi^{(\Delta)}$ is not orthonormal, and, moreover, it is linearly dependent. Tensor product of the bases $\psi^{(\Delta)} \otimes \chi$ forms orthonormal base of $\mathbb{T}^\pi \{ \Delta \}$, $\Delta \in \mathbb{N}$. In these tensor terms the embedding $\mathbb{T} \rightarrow \mathbb{T}^\pi$ is made by tensor product with the constant sequence $\psi^{(0,1)} \in \Pi$, $\psi_i^{(0,1)} \equiv 1$, $i \in \mathbb{Z}$: $\mathbb{T} \mapsto \mathbb{T} \otimes \psi^{(0,1)} = \mathbb{T}^\pi \{ 1 \} \subset \mathbb{T}^\pi$. Further for brevity we omit $\otimes \psi^{(0,1)}$ for subspaces $\mathbb{T} \subset \mathbb{T}$ and write $\mathbb{T} \subset \mathbb{T}^\pi$ instead of $\mathbb{T} \otimes \psi^{(0,1)} \subset \mathbb{T}^\pi$.

Similarly to the characters base (2) on $\mathbb{T}$, many important tangent subspaces $\mathbb{T} \subset \mathbb{T}^\pi$ are linear spans of subsets of $\psi^{(\Delta)} \otimes \chi$ for some $\Delta \in \mathbb{N}$. Therefore, combination of the characters base (2) and the tensor product representation (6) provides powerful tool for working in the tangent space $\mathbb{T}^\pi$. Below we illustrate it briefly.

Family of chi-square statistics «coded» by tangent subspaces

There is a family of chi-square statistics $S_{\mathbb{T}} = S_{\mathbb{T}} \left[x_1^n \right]$ bijectively «coded» by finite-dimensional tangent subspaces $\mathbb{T} \subset \mathbb{T}^\pi$. More precisely, for each $\mathbb{T}$ there is a class of asymptotically equivalent statistics as $n \rightarrow +\infty$. These equivalent statistics behave asymptotically identically for uniformly distributed binary sequence $\{ x_1^n \}$ and in some neighbourhood of this assumption called contigual asymptotics [4; 5], when $\ell \{ x \} \rightarrow u$ along the Markov manifold $\ell \{ x \} \in \mathbb{M}^\pi$ with a «speed» $O \left(n^{-\frac{1}{2}} \right)$ as $n \rightarrow +\infty$. Up to this equivalence, we can talk about statistics $S_{\mathbb{T}}$ without specifying which one is it from the equivalence class.

The family of statistics $S_{\mathbb{T}}$ can be derived from another family of $\mathbb{T}^\pi$ -valued normalised sufficient statistics (NSS) related to sufficient statistics of exponential families of Markov chains [6]:

$$W_{\mathbb{T}} = W_{\mathbb{T}} \left[x_1^n \right] \in \mathbb{T}, \quad S_{\mathbb{T}} = \| W_{\mathbb{T}} \|^2, \quad (11)$$

where norm is the Fisher – Riemann one. Similarly to $S_{\mathbb{T}}$, NSS $W_{\mathbb{T}}$ represents a class of asymptotically equivalent statistics as $n \rightarrow +\infty$. Under null hypothesis $W_{\mathbb{T}}$ has asymptotically standard normal distribution on $\mathbb{T} \subset \mathbb{T}^\pi$ with regard to the Fisher – Riemann dot product. Under contigual alternative H_1 (i. e. satisfying contigual asymptotics) limit normal distribution of $W_{\mathbb{T}}$ becomes non-centred but preserves identity covariance matrix. NSS (11) (i. e. their equivalence classes) are «projectively agreed»:

$$W_{T_2} = \text{proj} \left(W_{T_1} \middle| T_2 \right), \quad T_2 \subset T_1, \quad (12)$$

where $\text{proj}(\omega | \mathbb{T})$ is an orthogonal projection of $\omega \in \mathbb{T}^\pi$ onto finite-dimensional tangent subspace $\mathbb{T} \subset \mathbb{T}^\pi$ with regard to the Fisher – Riemann dot product.

Let us call T-test a chi-square statistical test of null hypothesis $H_0 : \ell \{ x \} = u$ associated with statistics $S_{\mathbb{T}}$. Under null hypothesis $S_{\mathbb{T}}$ is asymptotically chi-square distributed with $d = \dim(\mathbb{T})$ degrees of freedom: $\ell \{ S_{\mathbb{T}} | H_0 \} \rightarrow \chi_d^2$. Under contigual alternative $\ell \{ S_{\mathbb{T}} | H_1 \} \rightarrow \chi_{d, \lambda}^2$, where $\chi_{d, \lambda}^2$ is non-central chi-square distribution with non-centrality parameter $\lambda = \lim \| E \{ W_{\mathbb{T}} \} \|^2 \geq 0$, depending on space $\mathbb{T}$ and on deviation of alternative from null hypothesis (distance from $\ell \{ x \}$ to u in information metric). Asymptotic power of T-test at the significance level $\alpha \in (0, 1)$ equals $1 - F_{d, \lambda} \left(F_{d, 0}^{-1} (1 - \alpha) \right)$, where $F_{d, \lambda}(\cdot)$ and $F_{d, \lambda}^{-1}(\cdot)$ are cumulative distribution function and quantile function for $\chi_{d, \lambda}^2$, respectively. Table presents some tangent subspaces $\mathbb{T} \subset \mathbb{T}^\pi$ and their T-tests with equivalent analogues.

Let us fix some step $\Delta \in \mathbb{N}$, gram length $L \in \mathbb{N}$ and consider frequencies statistics

$$f^{(L, \Delta)}(q) = \sum_{i=0}^{N^{(L, \Delta)}-1} 1 \{ x_{i\Delta+j} = q_j, j \in [1:L] \}, \quad q \in \Upsilon^L, \quad N^{(L, \Delta)} = \left\lfloor \frac{n-L}{\Delta} \right\rfloor. \quad (13)$$

Denote standard chi-square statistics based on frequencies (13):

$$m^{(L, \Delta)} = \sum_{q \in Y^L} \frac{\left(f^{(L, \Delta)}(q) - 2^{-L} N^{(L, \Delta)} \right)^2}{2^{-L} N^{(L, \Delta)}}. \quad (14)$$

For $\Delta \geq L$ the L -grams in sum (13) do not overlap, whence $m^{(L, \Delta)}$ has asymptotic chi-square distribution under null hypothesis and belongs to the family $\{S_T\}_{T \subset T^\pi}$. For $\Delta < L$ this asymptotic distribution is generalised chi-square one and $m^{(L, \Delta)}$ does not belong to the family $\{S_T\}$. The following two types of chi-square statistics [1] based on formula (14) both belong to the family $\{S_T\}$:

$$M_1^{(s, \Delta)} = m^{((s+1)\Delta, \Delta)} - m^{(s\Delta, \Delta)}, \quad D_1^{(s, \Delta)} = 2^{s\Delta} (2^\Delta - 1), \quad s \geq 0, \quad (15)$$

$$M_2^{(s, \Delta)} = M_1^{(s, \Delta)} - M_1^{(s-1, \Delta)}, \quad D_2^{(s, \Delta)} = 2^{(s-1)\Delta} (2^\Delta - 1)^2, \quad s \geq 1, \quad (16)$$

where $m^{(0, \Delta)} \equiv 0$. Chi-square statistics $M_i^{(s, \Delta)}$ have $D_i^{(s, \Delta)}$ degrees of freedom. Statistics of the second type $M_2^{(s, \Delta)}$ are asymptotically mutually independent for different $s \geq 1$ under any fixed $\Delta \in \mathbb{N}$. T-test based on $M_1^{(0, \Delta)}$ is the standard test of uniform distribution of non-overlapping Δ -blocks $(x_{i\Delta+1}, x_{i\Delta+2}, \dots, x_{(i+1)\Delta})$ (FIPS Poker test [7] for $\Delta = 4$). Tests based on $M_1^{(s, 1)}$ and $M_2^{(s, 1)}$ are the NIST Serial tests [8] of types I and II, respectively (with parameter $m = s + 1$). Test based on $M_1^{(0, 1)}$ is the Monobit test. Test based on $M_2^{(1, 1)}$ is asymptotically equivalent to the Runs test.

A remarkable feature of the family of chi-square statistics $\{S_T\}_{T \subset T^\pi}$ following from properties (11) and (12) is that their asymptotic probabilistic properties are «consistent» with geometric properties of corresponding tangent subspaces $T \subset T^\pi$. Due to this feature, obtaining asymptotic probabilistic properties of $\{S_T\}$ comes down to calculation of geometric quantities based on Fisher – Riemann dot product or combinatorial quantities, e. g., for large orthonormal bases considered as combinatorial objects.

Introduce two operations on the subspaces $T_1, T_2 \subset T^\pi$: $T_1 \boxplus T_2$ for direct sum of orthogonal subspaces $T_1 \perp T_2$; $T_1 \boxminus T_2 := T_1 \cap T_2^\perp$ for orthogonal complement of nested subspace $T_2 \subset T_1$ within a larger one T_1 . The use of these operations implies that the spaces are obviously satisfy the properties of orthogonality (for $\boxplus$) or nesting (for $\boxminus$). For instance, $\partial T(s) = T(s) \boxminus T(s-1)$, $s \geq 1$, and $T = \lim_{s \rightarrow +\infty} T(s)$ breaks into a sum

$$T = \boxplus_{s \in \mathbb{N}_0} \partial T(s). \quad (17)$$

Tangent subspaces $T \subset T^\pi$ and their T-tests

Tangent subspace $T \subset T^\pi$	Degrees of freedom $\dim(T)$	Equivalent T-tests from NIST [8] and FIPS [7]
$T(0)$	1	NIST Monobit test, NIST Serial test (type I, $m = 1$)
$\partial T(1)$	1	NIST Runs test, NIST Serial test (type II, $m = 2$)
$T(s), s \geq 0$	2^s	NIST Serial test (type I, $m = s + 1$), NIST Approximate entropy test ($m = s + 1$)
$\partial T(s), s \geq 1$	2^{s-1}	NIST Serial test (type II, $m = s + 1$)
$\sum_{t \in \mathbb{Z}_4} T(t) \otimes \psi^{(t, 4)}$	15	FIPS Poker test

Theorem 1. *The following properties hold for subspaces $T_1, T_2 \subset T^\pi$:*

- under contiguous alternative S_{T_1} and S_{T_2} are asymptotically independent iff $T_1 \perp T_2$;
- $S_{T_1 \boxplus T_2} = S_{T_1} + S_{T_2}$, $S_{T_1 \boxminus T_2} = S_{T_1} - S_{T_2}$ (additivity).

Proof. Follows directly from properties of normal distribution, properties (11) and (12), and asymptotic normality of NSS $\{\mathbb{W}_T\}_{T \subset \mathbb{T}^\pi}$ with identity covariance matrices.

Proof is complete.

Let $\xi^{(i)} = \left(\xi_j^{(i)} \right)_{j=1}^{d_i} \in \mathbb{R}^{d_i}$, $i \in \{1, 2\}$, be two dependent standard Gaussian vectors with cross-covariations $\text{cov} \left\{ \xi_j^{(1)}, \xi_k^{(2)} \right\} = 1 \{j = k \leq d_3\} \theta_j$, $\theta_1, \dots, \theta_{d_3} \neq 0$, $d_3 \leq \min \{d_1, d_2\}$. Squared Euclidean norms of these Gaussian vectors are chi-square distributed: $\sigma_i = \left\| \xi^{(i)} \right\|^2 \sim \chi_{d_i}^2$. Joint probability distribution of $\sigma = (\sigma_1, \sigma_2) \in \mathbb{R}_+^2$ depends on (d_1, d_2) and on the multiset $\Lambda = \{\lambda_i\}_{i=1}^{d_3}$, $\lambda_i = \theta_i^2$. By multiset we mean possible multiplicity of its elements, e. g., multiset $\{1, 1, 2\}$ consists of 1 (twice) and 2 (once). Denote this bivariate chi-square distribution $\chi_{d_1, d_2}^{2 \times 2}(\Lambda) := \ell \{ \sigma \}$. The case of independence of σ_1 and σ_2 is $\Lambda = \emptyset$. The case of semi-independence with correlation coefficient $\text{cor} \{ \sigma_1, \sigma_2 \} = \frac{d_3}{d_1}$ is $d_1 = d_2$ and $\Lambda = \{1, \dots, 1\}$ consisting of d_3 ones. In general case

$$\text{cov} \{ \sigma_1, \sigma_2 \} = \sum_{j=1}^{d_3} \text{cov} \left\{ \left(\xi_j^{(1)} \right)^2, \left(\xi_j^{(2)} \right)^2 \right\} = 2 \sum_{\lambda \in \Lambda} \lambda. \quad (18)$$

Let $\text{proj}(T_1 | T_2)$, $T_1, T_2 \subset \mathbb{T}^\pi$, be linear operator $T_1 \rightarrow T_2$ of orthogonal projection of space T_1 onto T_2 ; $\Lambda(T_1, T_2)$ be a multiset of non-zero squared singular numbers of projection operator $\text{proj}(T_1 | T_2)$ (it is commutative $\Lambda(T_1, T_2) = \Lambda(T_2, T_1)$ as $\text{proj}(T_1 | T_2)$ and $\text{proj}(T_2 | T_1)$ are adjoint operators and have the same singular numbers). We further call $\Lambda(T_1, T_2)$ Jordan multiset of a pair of tangent subspaces because Jordan principal angles [9] are closely related to it.

Corollary 1. Let $T_1, T_2 \subset \mathbb{T}^\pi$ be arbitrary tangent subspaces of finite dimensions $d_i = \dim(T_i)$. Under null hypothesis the pair (S_{T_1}, S_{T_2}) is asymptotically $\chi_{d_1, d_2}^{2 \times 2}(\Lambda(T_1, T_2))$ -distributed, and its asymptotic covariance admits the following equivalent expressions:

$$\lim \text{cov} \{ S_{T_1}, S_{T_2} \} = 2 \sum_{\lambda \in \Lambda(T_1, T_2)} \lambda = 2 \sum_{j, k} \left\langle \omega^{(j)}, \tilde{\omega}^{(k)} \right\rangle^2 = 2 \left\| \text{proj}(T_1 | T_2) \right\|_{\text{Fr}}^2, \quad (19)$$

where $\{\omega^{(j)}\} \subset T_1$, $\{\tilde{\omega}^{(k)}\} \subset T_2$ are arbitrary orthonormal bases; $\|\cdot\|_{\text{Fr}}$ is a Frobenius norm of linear operator.

Proof. Projection operator $\text{proj}(T_1 | T_2)$ is represented by a matrix of entries $\left\langle \omega^{(j)}, \tilde{\omega}^{(k)} \right\rangle$ in a pair of bases $\{\omega^{(j)}\}$, $\{\tilde{\omega}^{(k)}\}$. Sum of these squared entries and sum of squared singular numbers are two representations of squared Frobenius norm for $\text{proj}(T_1 | T_2)$, whence we get the second and the third equalities of expression (19). Now let us choose the left and the right singular vectors of $\text{proj}(T_1 | T_2)$ for the bases $\{\omega^{(j)}\}$, $\{\tilde{\omega}^{(k)}\}$, and set asymptotically standard normal vectors:

$$\xi^{(1)} = \left(\xi_j^{(1)} \right), \quad \xi^{(2)} = \left(\xi_k^{(2)} \right), \quad \xi_j^{(1)} = \left\langle \mathbb{W}_T, \omega^{(j)} \right\rangle, \quad \xi_k^{(2)} = \left\langle \mathbb{W}_T, \tilde{\omega}^{(k)} \right\rangle, \quad T = T_1 + T_2.$$

Their asymptotic cross-covariations by construction are

$$\lim \text{cov} \left\{ \xi_j^{(1)}, \xi_k^{(2)} \right\} = \left\langle \omega^{(j)}, \tilde{\omega}^{(k)} \right\rangle = 1 \{j = k \leq d_3\} \sqrt{\lambda_j}, \quad \Lambda(T_1, T_2) = \{\lambda_j\}_{j=1}^{d_3}.$$

By properties (11) and (12), we have $\left\| \xi^{(r)} \right\|^2 = S_{T_r}$, $r \in \{1, 2\}$, whence follows the needed asymptotic distribution of the pair (S_{T_1}, S_{T_2}) according to definition of $\chi^{2 \times 2}$ -distribution. The first equality of expression (19) follows from formula (18).

Proof is complete.

Let us call subspaces $T_1, T_2 \subset \mathbb{T}^\pi$ semi-orthogonal if $T_1 \boxplus T_3 \perp T_2 \boxplus T_3, T_3 = T_1 \cap T_2$. Nested spaces ($T_2 \subset T_1$) and orthogonal ones ($T_1 \perp T_2$) are the special cases. We also call two identically chi-square distributed random variables $\sigma_1, \sigma_2 \sim \chi_D^2$ semi-independent [1] if they are representable as

$$\sigma_1 = \tilde{\sigma}_1 + \tilde{\sigma}_0, \quad \sigma_2 = \tilde{\sigma}_2 + \tilde{\sigma}_0, \quad \tilde{\sigma}_0 \sim \chi_d^2, \quad \tilde{\sigma}_1, \tilde{\sigma}_2 \sim \chi_{D-d}^2, \quad (20)$$

for some $0 \leq d \leq D$ and mutually independent $\{\tilde{\sigma}_i\}_{i=0}^2$. Obviously $\text{cor}\{\sigma_1, \sigma_2\} = \frac{d}{D}$.

Corollary 2. Let $T_1, T_2 \subset \mathbb{T}^\pi$ be subspaces of equal dimensions $\dim(T_1) = \dim(T_2)$. Under null hypothesis statistics S_{T_1} and S_{T_2} are asymptotically semi-independent with the asymptotic correlation coefficient

$$\lim \text{cor}\{S_{T_1}, S_{T_2}\} = \frac{\dim(T_1 \cap T_2)}{\dim(T_1)}$$

iff their spaces T_1, T_2 are semi-orthogonal.

Proof. Corollary 1 states bijection between asymptotic joint distribution of a pair (S_{T_1}, S_{T_2}) and relative position of spaces T_1, T_2 . In this bijection representation of the form (20) corresponds to the following representation for spaces:

$$T_1 = \tilde{T}_1 \boxplus \tilde{T}_0, \quad T_2 = \tilde{T}_2 \boxplus \tilde{T}_0, \quad \tilde{T}_1 \perp \tilde{T}_2,$$

which is equivalent to semi-orthogonality of T_1, T_2 .

Proof is complete.

Consider a class of tangent subspaces $T \subset \mathbb{T}^\pi$ to illustrate the above theory. Let us fix some period $\Delta \in \mathbb{N}$. Due to sum (17), the space $\mathbb{T}^\pi\{\Delta\} = \mathbb{T} \otimes \Pi\{\Delta\}$ breaks into a sum

$$\mathbb{T}^\pi\{\Delta\} = \boxplus_{(t,h) \in C_\Delta} \psi^{(t,\Delta)} \otimes \partial T(h), \quad C_\Delta := \mathbb{Z}_\Delta \times \mathbb{N}_0. \quad (21)$$

We call Δ -periodic cell of spatio-temporal index (t, h) (index for brevity, with temporal part t and spatial part h) the summand subspace under $\boxplus$ in formula (21), and we call Δ -periodic cellular space of profile $C \subset C_\Delta, |C| < +\infty$, a partial sum of Δ -periodic cells with indices from the profile C :

$$\mathbb{T}^\pi\{C, \Delta\} := \boxplus_{(t,h) \in C} \psi^{(t,\Delta)} \otimes \partial T(h), \quad \dim(\mathbb{T}^\pi\{C, \Delta\}) =: \dim(C) = \sum_{(t,h) \in C} \gamma[h], \quad (22)$$

where $\dim(\cdot)$ is an unbounded integer-valued dimensionality measure on C_Δ with atomic weights $\gamma[h]$ determined by formula (4).

Example 1. In particular, chi-square statistics (15) and (16) correspond to cellular spaces. Let us show how to get their profiles $C \subset C_\Delta$ by example of FIPS Poker test [7] that corresponds to statistics (15) with $\Delta = 4$ and $s = 0$. The expression in table for tangent subspace $T_{\text{Poker}} \subset \mathbb{T}^\pi$ of Poker test is reduced to the form (22) by breaking $\mathbb{T}(t)$ into sum $\boxplus_{0 \leq \tau \leq t} \partial T(\tau)$. The expression from table, in its turn, is obtained as follows. Space T_{Poker} is a tangent space of u within 15-dimensional Markov model of independent and identically distributed binary 4-blocks. Without loss of generality 4-block starts at zero time: $(x_0, \dots, x_3) \in Y^4$. Let us represent this Markov model as a 4-periodic one of the form (5). Symbol x_0 does not depend on prehistory $\{x_i : i < 0\}$, whence $\omega_0 \in \mathbb{T}(0)$ (zero Markov order). Symbol x_1 under its fixed prehistory $\{x_i : i < 1\}$ conditionally depends only on x_0 , whence $\omega_1 \in \mathbb{T}(1)$ (first Markov order). Similarly $\omega_i \in \mathbb{T}(i), i \in \mathbb{Z}_4$. These conditions are equivalent to representation in table. So the profile of 4-periodic cellular space T_{Poker} is $C = \{(t, h) : 0 \leq h \leq t \in \mathbb{Z}_4\}$, and similarly for statistics (15) with any arbitrary $\Delta \in \mathbb{N}$ and $s = 0$ («stair» shape on fig. 1). Profiles corresponding to statistics (16) have «double-stair» shapes (fig. 2) obtained by set subtraction of two «stair» shapes of different «heights».

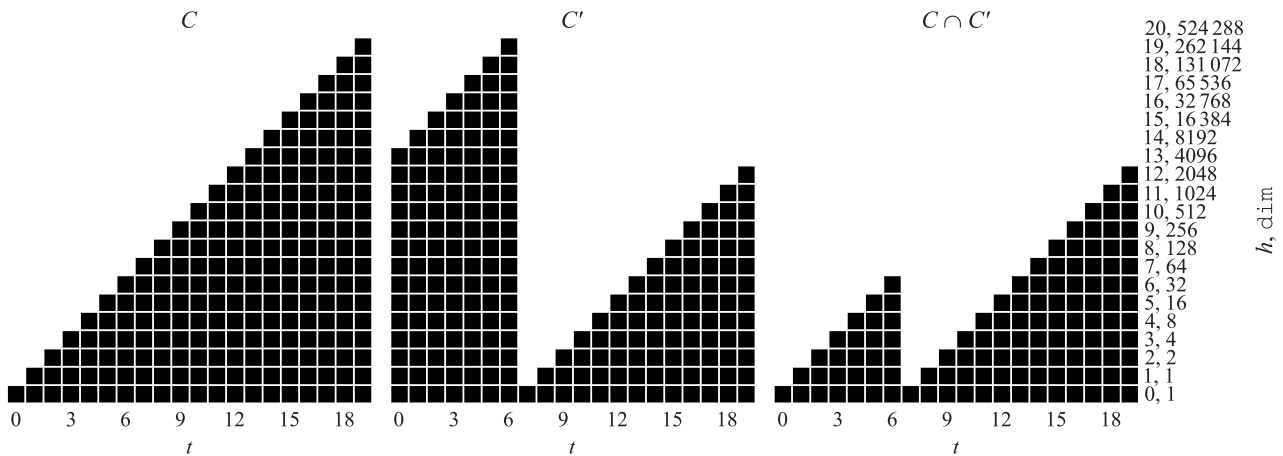


Fig. 1. Illustration to expression (23) for «stair» profile C and its cyclic shift C' ($\Delta = 20$, $\delta = 7$)

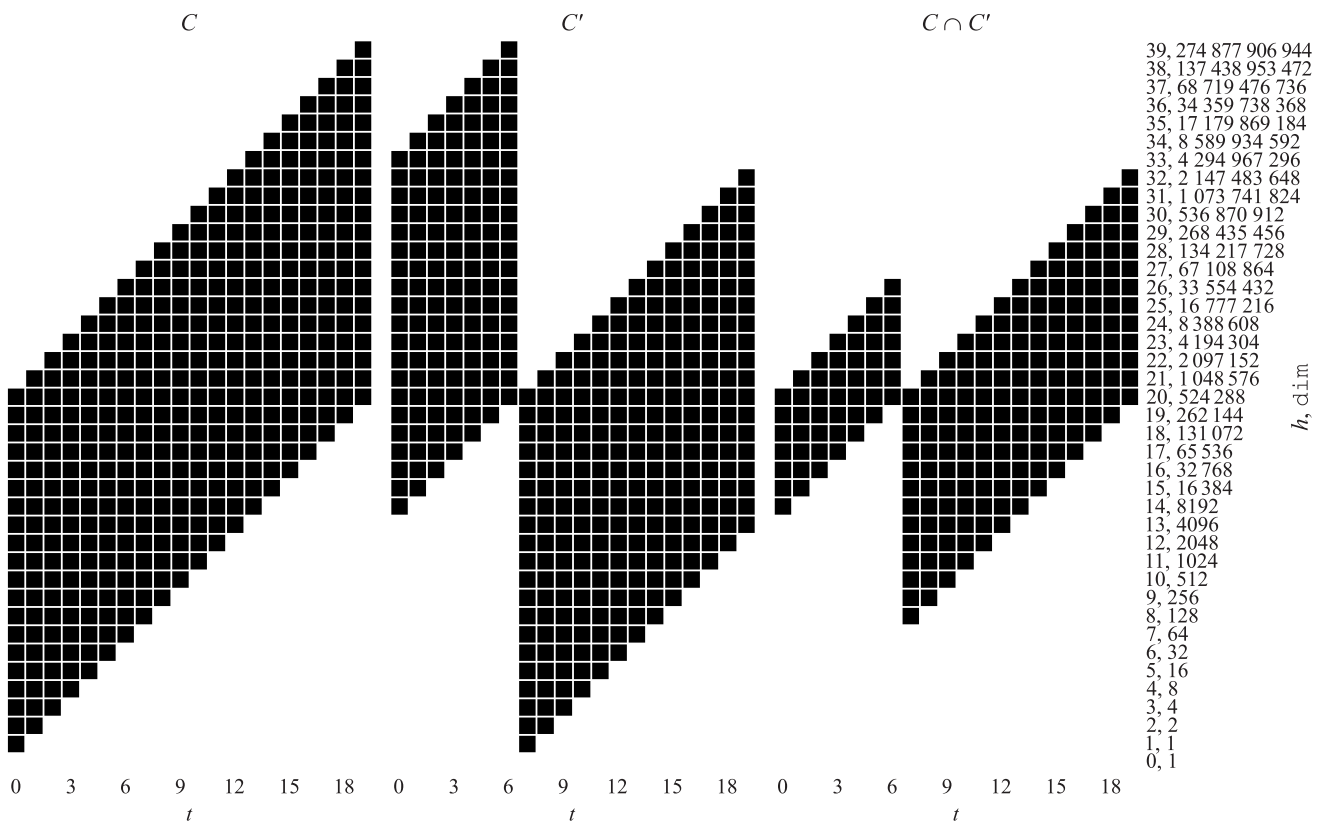


Fig. 2. Illustration to expression (23) for «double-stair» profile C and its cyclic shift C' ($\Delta = 20$, $\delta = 7$)

Due to corollary 2, asymptotic semi-independence of chi-square statistics is equivalent to semi-orthogonality of their tangent subspaces. Let us fix some $\Delta \in \mathbb{N}$ and two profiles $C, C' \subset C_\Delta$ of equal dimensionality measure $d = \dim(C) = \dim(C')$. Corresponding Δ -periodic cellular spaces of equal dimensions d are semi-orthogonal by definition, whence chi-square statistics $S_{\mathbb{T}^\pi\{C, \Delta\}}, S_{\mathbb{T}^\pi\{C', \Delta\}}$ are asymptotically semi-independent, and finding their asymptotic correlation coefficient $\frac{\tilde{d}}{d}$ comes down to calculation of dimensionality measure of intersection profile:

$$\tilde{d} = \dim(\mathbb{T}^\pi\{C, \Delta\} \cap \mathbb{T}^\pi\{C', \Delta\}) = \dim(\mathbb{T}^\pi\{C \cap C', \Delta\}) = \dim(C \cap C').$$

Let us fix now some profile $C \subset C_\Delta$ and consider corresponding chi-square statistics in two versions: undistorted $S_{\mathbb{T}^\pi\{C, \Delta\}}[\mathbf{x}_1^n]$ and δ -distorted $S_{\mathbb{T}^\pi\{C, \Delta\}}[\mathbf{x}_{1+\delta}^n]$ (without the first δ «bits» (see [1, sect. 4])). Reasoning

similarly to example 1, it is easy to show that δ -distortion leads to cyclic δ -shifting of profile along the temporal axis:

$$S_{\mathbb{T}^\pi\{C, \Delta\}} \left[\mathbf{x}_{1+\delta}^n \right] = S_{\mathbb{T}^\pi\{C', \Delta\}} \left[\mathbf{x}_1^n \right], \quad C' = \mathfrak{R}^\delta C,$$

where $\mathfrak{R}^\delta: \mathbf{C}_\Delta \rightarrow \mathbf{C}_\Delta$, $\mathfrak{R}^\delta(t, h) := \left(t + \frac{\Delta}{2}, h \right)$, $(t, h) \in \mathbf{C}_\Delta$. Obviously, dimensionality measure is invariant under this cyclic shift $\dim(C) = \dim(\mathfrak{R}^\delta C)$ as its weight $\gamma[h]$ depends only on the spatial part h of index (t, h) . Finally for chi-square statistics corresponding to Δ -periodic cellular spaces the asymptotic correlation coefficient between the undistorted and δ -distorted versions admits the following expression:

$$\limcor \left\{ S_{\mathbb{T}} \left[\mathbf{x}_1^n \right], S_{\mathbb{T}} \left[\mathbf{x}_{1+\delta}^n \right] \right\} = \frac{\dim(C \cap \mathfrak{R}^\delta C)}{\dim(C)}, \quad \mathbb{T} = \mathbb{T}^\pi\{C, \Delta\}. \quad (23)$$

Below we prove the result presented in article [1] using relation (23) for «stair» and «double-stair» profiles $C \subset \mathbf{C}_\Delta$ (see example 1). For these two types of profiles fig. 1 and 2 illustrate components of the right-hand side fraction (23): original profile C , cyclically shifted one $\mathfrak{R}^\delta C$, and their intersection $C \cap \mathfrak{R}^\delta C$. Weights $\gamma[h]$ of dimensionality measure are plotted on the vertical axis: their sum over profile (black cells) is equal to dimensionality of corresponding cellular space.

Theorem 2. For $0 \leq \delta \leq \Delta$ undistorted and δ -distorted statistics (15) and (16) are asymptotically semi-independent with the following asymptotic correlation coefficients under null hypothesis:

$$\limcor \left\{ M_1^{(s, \Delta)} \left[\mathbf{x}_1^n \right], M_1^{(s, \Delta)} \left[\mathbf{x}_{1+\delta}^n \right] \right\} = \frac{2^\delta + 2^{\delta'} - 2}{2^\Delta - 1}, \quad (24)$$

$$\limcor \left\{ M_2^{(s, \Delta)} \left[\mathbf{x}_1^n \right], M_2^{(s, \Delta)} \left[\mathbf{x}_{1+\delta}^n \right] \right\} = \frac{(2^\delta + 2^{\delta'})(2^\Delta + 1) - 2^{2+\Delta}}{(2^\Delta - 1)^2}, \quad (25)$$

where $\delta' = \Delta - \delta$.

Proof. «Stair» profile for $M_1^{(s, \Delta)}$ is

$$C = C_1^{(s, \Delta)} = \{(t, h) : 0 \leq h \leq s\Delta + t, t \in \mathbb{Z}_\Delta\}. \quad (26)$$

Corresponding δ -shifted profile in expression (23) is

$$C' = \{(t, h) : 0 \leq h \leq s\Delta + \tilde{t}, t \in \mathbb{Z}_\Delta\}, \quad \tilde{t} = \left(t + \frac{\Delta}{2} \right).$$

The following sums are used below:

$$\begin{aligned} \sum_{t \in \mathbb{Z}_\Delta} 2^{\min\{t, \tilde{t}\}} &= \sum_{\tau=0}^{\delta-1} 2^\tau + \sum_{\tau=0}^{\delta'-1} 2^\tau = 2^\delta + 2^{\delta'} - 2, \\ \sum_{t \in \mathbb{Z}_\Delta} 2^{\max\{t, \tilde{t}\}} &= \sum_{t \in \mathbb{Z}_\Delta} \left(2^t + 2^{\tilde{t}} - 2^{\min\{t, \tilde{t}\}} \right) = 2^{1+\Delta} - 2^\delta - 2^{\delta'}. \end{aligned} \quad (27)$$

Using formulas (4) and (27), dimensionality measure for intersection

$$\dim(C \cap C') = \sum_{t \in \mathbb{Z}_\Delta} \sum_{h=0}^{s\Delta + \min\{t, \tilde{t}\}} \gamma[h] = 2^{s\Delta} \sum_{t \in \mathbb{Z}_\Delta} 2^{\min\{t, \tilde{t}\}} = 2^{s\Delta} (2^\delta + 2^{\delta'} - 2).$$

We put this and $\dim(C) = D_1^{(s, \Delta)} = 2^{s\Delta} (2^\Delta - 1)$ into expression (23) and get asymptotic correlation coefficient (25).

«Double-stair» profile for $M_2^{(s, \Delta)}$ is a set subtraction of two adjacent «stair» profiles (26):

$$C = C_2^{(s, \Delta)} = C_1^{(s, \Delta)} \setminus C_1^{(s-1, \Delta)} = \{(t, h) : (s-1)\Delta + t < h \leq s\Delta + t, t \in \mathbb{Z}_\Delta\}.$$

Corresponding δ -shifted profile is

$$C' = \{(t, h) : (s-1)\Delta + \tilde{t} < h \leq s\Delta + \tilde{t}, t \in \mathbb{Z}_\Delta\}.$$

Using formula (27), dimensionality measure for intersection

$$\begin{aligned} \dim(C \cap C') &= \sum_{t \in \mathbb{Z}_\Delta} \sum_{h=1+(s-1)\Delta+\max\{t, \tilde{t}\}}^{s\Delta+\min\{t, \tilde{t}\}} \gamma[h] = \sum_{t \in \mathbb{Z}_\Delta} \left(2^{s\Delta+\min\{t, \tilde{t}\}} - 2^{(s-1)\Delta+\max\{t, \tilde{t}\}} \right) = \\ &= 2^{s\Delta} (2^\delta + 2^{\delta'} - 2) - 2^{(s-1)\Delta} (2^{1+\Delta} - 2^\delta - 2^{\delta'}). \end{aligned}$$

Put this and $\dim(C) = D_2^{(s, \Delta)} = 2^{(s-1)\Delta} (2^\Delta - 1)^2$ into expression (23) and get asymptotic correlation coefficient (25) after simplification.

Proof is complete.

Note that the values (24) and (25) do not depend on parameter s . Their plots for $\Delta = 7$ are presented in fig. 3. These plots are very close to each other, and they indeed are asymptotically equivalent (i. e. ratio goes to unity) uniformly over $0 \leq \delta \leq \Delta$ as $\Delta \rightarrow +\infty$. The plots are minimal at half-period $\delta = \frac{\Delta}{2}$ ($\delta = \frac{\Delta \pm 1}{2}$ for odd Δ) and maximal at full period $\delta \in \{0, \Delta\}$. The minimal values have asymptotics $\text{cor}_{\min} \sim R \cdot 2^{-\frac{\Delta}{2}}$, where $R = 2$ for even Δ and $R = \frac{3\sqrt{2}}{2}$ for odd Δ . In the neighbourhood of half-period the plots tend to the scaled hyperbolic cosine. In the neighbourhood of full period adjacent values differ approximately twofold for large Δ .

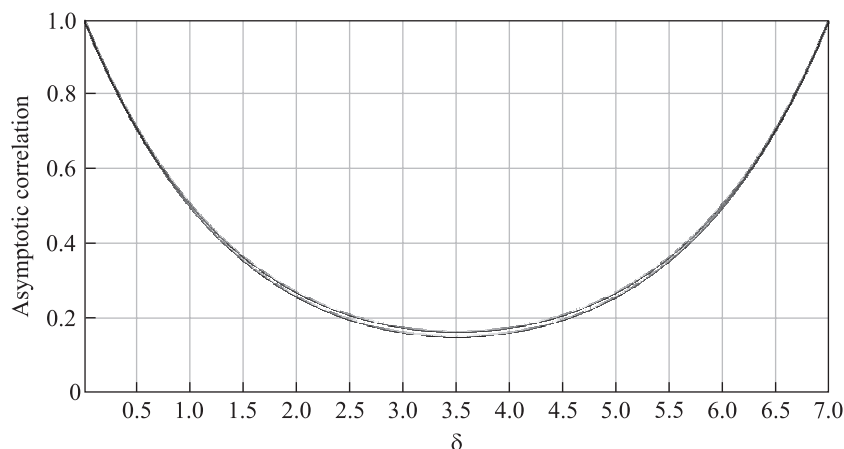


Fig. 3. Asymptotic correlation coefficients (24) and (25) plotted against δ ($\Delta = 7$)

Consider now joint asymptotic distribution of a pair $\left(S_{\mathbb{T}^\pi\{C, \Delta\}}, S_{\mathbb{T}^\pi\{C', \Delta'\}} \right)$ of chi-square statistics corresponding to cellular spaces of two arbitrary periods $\Delta, \Delta' \in \mathbb{N}$. In the case $\Delta \neq \Delta'$ there is no more guaranteed semi-orthogonality of cellular spaces, and as a consequence, asymptotic semi-independence of chi-square statistics. Therefore, we need more general corollary 1 to describe joint asymptotic distribution of chi-square statistics. The following trivial factorisation lemma is also used.

Lemma 1. Let two sums $T_1 = \bigoplus_{j \in J} T_{1,j}$, $T_2 = \bigoplus_{k \in K} T_{2,k}$ of tangent subspaces satisfy diagonality property $T_{1,j} \perp T_{2,k}$ for any $j \neq k$. Then their Jordan multiset breaks into union

$$\Lambda(T_1, T_2) = \bigcup_{j \in J \cap K} \Lambda(T_{1,j}, T_{2,j}).$$

Let us break $\gamma[h]$ -dimensional Δ -periodic cell of index $(t, h) \in C_\Delta$ into a sum of one-dimensional orthogonal subspaces:

$$\Psi^{(t, \Delta)} \otimes \partial \mathbb{T}(h) = \bigoplus_{Q \in \Upsilon^+[h]} \text{span} \left\{ \Psi^{(t, \Delta)} \otimes \chi^{(Q)} \right\}. \quad (28)$$

By analogy with Δ -periodic cells, we call Δ -periodic atom of spatio-temporal index (index for brevity) $(t, Q) \in \mathbf{A}_\Delta := \mathbb{Z}_\Delta \times \Upsilon^+$ the summand subspace under $\bigoplus$ in formula (28): Q is a spatial part of atom's index. Rewrite cellular space (22) as a sum of atoms, grouping summands by spatial parts Q :

$$\mathbb{T}^\pi\{C, \Delta\} = \bigoplus_{(t, h) \in C} \bigoplus_{Q \in \mathbb{Y}^+[h]} \text{span}\{\psi^{(t, \Delta)} \otimes \chi^{(Q)}\} = \bigoplus_{h \in \mathbf{h}^{(C)}} \bigoplus_{Q \in \mathbb{Y}^+[h]} \chi^{(Q)} \otimes \Pi\{\mathbf{t}_h^{(C)}, \Delta\}, \quad (29)$$

$$\Pi\{\mathbf{t}, \Delta\} := \text{span}\{\psi^{(t, \Delta)} : t \in \mathbf{t}\} \subset \Pi\{\Delta\},$$

$$\mathbf{t}_h^{(C)} := \{t : (t, h) \in C\} \subset \mathbb{Z}_\Delta, \quad \mathbf{h}^{(C)} := \{h : \exists (t, h) \in C\} \subset \mathbb{N}_0. \quad (30)$$

Here $\mathbf{t}_h^{(C)}$ is a «temporal slice» of profile C at the spacial level h ; $\mathbf{h}^{(C)}$ is a «spatial projection» of profile C ; $\Pi\{\mathbf{t}, \Delta\}$ is a canonic subspace of $\Pi\{\Delta\}$ corresponding to «temporal profile» $\mathbf{t}$.

Theorem 3. *Jordan multiset of two arbitrary cellular spaces is as follows:*

$$\Lambda(\mathbb{T}^\pi\{C, \Delta\}, \mathbb{T}^\pi\{C', \Delta'\}) = \left\{ \left(\frac{|\mathbf{t}_{h, \tau}^{(C)}| |\mathbf{t}_{h, \tau}^{(C')}|}{aa'} \right)^{\times \gamma[h]} : h \in \mathbf{h}^{(C)} \cap \mathbf{h}^{(C')}, \tau \in \left(\frac{\mathbf{t}_h^{(C)}}{\kappa \mathbb{Z}} \right) \cap \left(\frac{\mathbf{t}_h^{(C')}}{\kappa \mathbb{Z}} \right) \right\}, \quad (31)$$

where notation (30) is used, and also $\lambda^{\times \mu}$ means element λ of multiset taken with multiplicity μ ; $\kappa = \gcd(\Delta, \Delta')$; $a = \frac{\Delta}{\kappa}$; $a' = \frac{\Delta'}{\kappa}$; $\mathbf{t}_{h, \tau}^{(C)} = \left\{ t \in \mathbf{t}_h^{(C)} : t \equiv \tau \right\}$.

Proof. Consider a similar sum (29) for cellular space $\mathbb{T}^\pi\{C', \Delta'\}$. Diagonality property of lemma 1 holds for these sums over Q : it follows from formula (9) and characters orthogonality. Hence

$$\begin{aligned} \Lambda(\mathbb{T}^\pi\{C, \Delta\}, \mathbb{T}^\pi\{C', \Delta'\}) &= \bigcup_{h \in \mathbf{h}^{(C)} \cap \mathbf{h}^{(C')}, Q \in \mathbb{Y}^+[h]} \Lambda(\chi^{(Q)} \otimes \Pi\{\mathbf{t}_h^{(C)}, \Delta\}, \chi^{(Q)} \otimes \Pi\{\mathbf{t}_h^{(C')}, \Delta'\}) = \\ &= \bigcup_{h \in \mathbf{h}^{(C)} \cap \mathbf{h}^{(C')}} \Lambda^{\times \gamma[h]}(\Pi\{\mathbf{t}_h^{(C)}, \Delta\}, \Pi\{\mathbf{t}_h^{(C')}, \Delta'\}), \end{aligned} \quad (32)$$

where $\Lambda^{\times \mu}$ means multiset Λ taken with multiplicity μ . Denote $\mathbf{t} = \mathbf{t}_h^{(C)}$, $\mathbf{t}' = \mathbf{t}_h^{(C')}$ for brevity. From formulas (8) and (10) we get

$$\langle \psi^{(t, \Delta)}, \psi^{(t', \Delta')} \rangle = 1 \left\{ t \equiv t' \right\} (aa')^{-\frac{1}{2}}, \quad t \in \mathbb{Z}_\Delta, \quad t' \in \mathbb{Z}_{\Delta'}. \quad (33)$$

In other words, $\langle \psi^{(t, \Delta)}, \psi^{(t', \Delta')} \rangle \equiv 0$ for $t \not\equiv t'$. Let us break $\Pi\{\mathbf{t}, \Delta\}$ into a sum over integers modulo κ :

$$\Pi\{\mathbf{t}, \Delta\} = \bigoplus_{\tau \in \frac{\mathbf{t}}{\kappa \mathbb{Z}}} \Pi\{\mathbf{t}_\tau, \Delta\}, \quad \mathbf{t}_\tau = \left\{ t \in \mathbf{t} : t \equiv \tau \right\},$$

and similarly for $\Pi\{\mathbf{t}', \Delta'\}$. These sums satisfy diagonality property of lemma, whence

$$\Lambda(\Pi\{\mathbf{t}, \Delta\}, \Pi\{\mathbf{t}', \Delta'\}) = \bigcup_{\tau \in \left(\frac{\mathbf{t}}{\kappa \mathbb{Z}} \right) \cap \left(\frac{\mathbf{t}'}{\kappa \mathbb{Z}} \right)} \Lambda(\Pi\{\mathbf{t}_\tau, \Delta\}, \Pi\{\mathbf{t}'_\tau, \Delta'\}). \quad (34)$$

Due to formula (33), projection operator $\text{proj}(\Pi\{\mathbf{t}_\tau, \Delta\} | \Pi\{\mathbf{t}'_\tau, \Delta'\})$ is represented by a matrix $|\mathbf{t}_\tau| \times |\mathbf{t}'_\tau|$ of all equal entries $(aa')^{-\frac{1}{2}}$. This matrix has a single non-zero singular number $\sqrt{\frac{|\mathbf{t}_\tau| |\mathbf{t}'_\tau|}{aa'}}$, whence formula (34) takes the form

$$\Lambda(\Pi\{\mathbf{t}, \Delta\}, \Pi\{\mathbf{t}', \Delta'\}) = \left\{ \frac{|\mathbf{t}_\tau| |\mathbf{t}'_\tau|}{aa'} : \tau \in \left(\frac{\mathbf{t}}{\kappa \mathbb{Z}} \right) \cap \left(\frac{\mathbf{t}'}{\kappa \mathbb{Z}} \right) \right\}. \quad (35)$$

Combining formulas (32) and (35), we get multiset (31).

Proof is complete.

Theorem 3 means, in particular, that Jordan multiset of any two cellular spaces of periods $\Delta, \Delta' \in \mathbb{N}$ contains only rational numbers of the form $\frac{bb'}{aa'}$, $1 \leq b \leq a, 1 \leq b' \leq a'$, where $\frac{a}{a'}$ is an irreducible fraction equals to $\frac{\Delta}{\Delta'}$.

From theorem 3 we also derive the following semi-orthogonality criterion for cellular spaces (i. e. criterion for Jordan multiset of all units).

Corollary 3. *In the notation of theorem 3 cellular spaces $\mathbb{T}^\pi\{C, \Delta\}$ and $\mathbb{T}^\pi\{C', \Delta'\}$ are semi-orthogonal iff for each simultaneously non-empty equivalence classes*

$$\mathbf{t}_{h,\tau}^{(C)}, \mathbf{t}_{h,\tau}^{(C')} \neq \emptyset, h \in \mathbf{h}^{(C)} \cap \mathbf{h}^{(C')}, \tau \in \left(\frac{\mathbf{t}_h^{(C)}}{\kappa\mathbb{Z}} \right) \cap \left(\frac{\mathbf{t}_h^{(C')}}{\kappa\mathbb{Z}} \right),$$

these classes are both maximal:

$$|\mathbf{t}_{h,\tau}^{(C)}| = a, |\mathbf{t}_{h,\tau}^{(C')}| = a'.$$

This criterion obviously holds for $\Delta = \Delta'$ as $a = a' = 1$ and non-emptiness is equivalent to maximality. Introduce piecewise linear functions [1], satisfying finite difference relation (fig. 4):

$$\rho_1(x) = \max\{0, 1 - \max\{0, x\}\}, \rho_2(x) = \max\{0, 1 - |x|\} = \rho_1(x) - \rho_1(x+1), x \in \mathbb{R}. \quad (36)$$

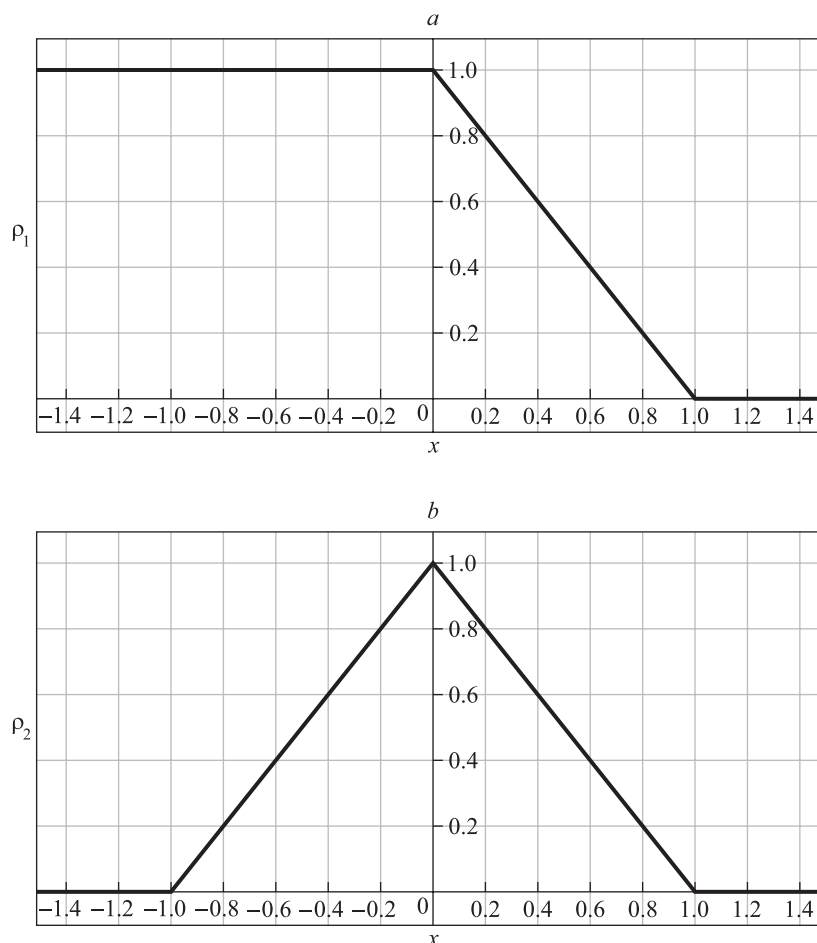


Fig. 4. Functions $\rho_1(x)$ (a) and $\rho_2(x)$ (b)

Below we prove the result presented in article [1], applying theorem 3 to «stair» and «double-stair» profiles.

Theorem 4. *Let $(i, j) \in \{(2, 2), (1, 2), (1, 1)\}$. In the notation of theorem 3 and functions (36) asymptotic covariances of statistics (15) and (16) under null hypothesis are as follows:*

$$\lim_{\text{cov}} \left\{ M_i^{(s, \Delta)}, M_j^{(s', \Delta')} \right\} = R + 2 \frac{(2^\kappa - 1)^2}{2^\kappa} \sum_{\mu=\mu_-}^{\mu_+} \rho_i \left(\frac{\mu}{a} - s \right) \rho_j \left(\frac{\mu}{a'} - s' \right) 2^{\mu\kappa}, \quad (37)$$

where $R = 2(2^\kappa - 1)$ for $(i, j) = (1, 1)$ and $R = 0$ otherwise; $\mu_+ = \min \{(s+1)a, (s'+1)a'\} - 1$;

$$\mu_- = \begin{cases} \max \{(s-1)a, (s'-1)a'\} + 1, & (i, j) = (2, 2), \\ (s'-1)a' + 1, & (i, j) = (1, 2), \\ 1, & (i, j) = (1, 1). \end{cases}$$

Proof. The cases $(i, j) = (2, 2)$ and $(i, j) = (1, 2)$ follow from the case $(i, j) = (1, 1)$, finite difference relations (16) and (36), and linearity of covariance with regard to its arguments. Thus, it is enough to prove asymptotic covariances (37) for $(i, j) = (1, 1)$, which, due to expression (19), is reduced to the summation of the elements over Jordan multiset (31) for «stair» profiles $C = C_1^{(s, \Delta)}$, $C' = C_1^{(s', \Delta')}$ of the form (26).

Let us fix some $\tau \in \mathbb{Z}_\kappa$ and find a part of sum corresponding to it (removing restrictions from the range of «stair heights» h adds zeros to Λ_τ without affecting the resulting sum):

$$\Sigma_\tau = \sum_{\lambda \in \Lambda_\tau} \lambda, \quad \Lambda_\tau = \left\{ \left(\frac{|\mathbf{t}_{h,\tau}^{(C)}| |\mathbf{t}_{h,\tau}^{(C')}|}{aa'} \right)^{\gamma[h]} : h \in \mathbb{N}_0 \right\}. \quad (38)$$

Break $\mathbb{N}_0$ into disjunctive intervals $I_\mu = [\max \{0, (\mu-1)\kappa + \tau + 1\} : \mu\kappa + \tau]$, $\mu \in \mathbb{N}_0$. For $h \leq s\Delta$ we have $\mathbf{t}_h^{(C)} = [0 : \Delta - 1]$. For $s\Delta < h < (s+1)\Delta$, $\mathbf{t}_h^{(C)} = [h - s\Delta : \Delta - 1]$. For $h \geq (s+1)\Delta$, $\mathbf{t}_h^{(C)} = \emptyset$. Finally it is easy to check the following relation for the number $|\mathbf{t}_{h,\tau}^{(C)}|$ of elements in $\mathbf{t}_h^{(C)}$ that equal τ modulo κ :

$$\frac{|\mathbf{t}_{h,\tau}^{(C)}|}{a} = \rho_1 \left(\frac{\mu}{a} - s \right), \quad h \in I_\mu, \quad \mu \in \mathbb{N}_0. \quad (39)$$

Putting expression (39) and similar relation for C' into formula (38) leads to the following expression (formula (4) is used):

$$\begin{aligned} \Sigma_\tau &= \sum_{\mu \in \mathbb{N}_0} \sum_{h \in I_\mu} \gamma[h] \frac{|\mathbf{t}_{h,\tau}^{(C)}| |\mathbf{t}_{h,\tau}^{(C')}|}{aa'} = \sum_{\mu \in \mathbb{N}_0} \rho_1 \left(\frac{\mu}{a} - s \right) \rho_1 \left(\frac{\mu}{a'} - s' \right) \sum_{h \in I_\mu} \gamma[h] = \\ &= \sum_{h \in I_0} \gamma[h] + \sum_{\mu \in \mathbb{N}} \rho_1 \left(\frac{\mu}{a} - s \right) \rho_1 \left(\frac{\mu}{a'} - s' \right) \sum_{h \in I_\mu} \gamma[h] = 2^\tau + 2^{\tau-\kappa} (2^\kappa - 1) \sum_{\mu \in \mathbb{N}} 2^{\mu\kappa} \rho_1 \left(\frac{\mu}{a} - s \right) \rho_1 \left(\frac{\mu}{a'} - s' \right). \end{aligned} \quad (40)$$

Summation of expressions (40) over $\tau \in \mathbb{Z}_\kappa$ with doubling (see expression (19)) gives asymptotic covariances (37) for $(i, j) = (1, 1)$, where the upper limit μ_+ of summation over μ «cuts off» zero summands at $\mu > \mu_+$.

Proof is complete.

Thus, asymptotic covariances of statistics (15) and (16) are expressed as some weighted dot products over integer lattice of «shifts and stretchings» of functions (36).

References

1. Voloshko VA. Local information geometry for high-order binary Markov chains and its applications. In: Kharin YuS, Zubkov AM, Kharin AYU, Maltsev MV, Palukha UYu, editors. *Computer data analysis and modeling: stochastics and data science. Proceedings of the 14th International conference; 2025 September 24–27; Minsk, Belarus*. Minsk: Belarusian State University; 2025. p. 266–271.
2. Amari S, Nagaoka H. *Methods of information geometry*. Harada D, translator. Oxford: Oxford University Press; 2000. 206 p. (Translations of mathematical monographs; volume 191).

3. Luong B. *Fourier analysis on finite Abelian groups*. Boston: Birkhauser; 2009. XVI, 159 p. (Applied and numerical harmonic analysis).
4. Voloshko VA, Kharin YuS, Trubey AI. On power comparison for some tests on pure randomness under Markov high-order dependencies. In: Kharin YuS, Zubkov AM, Kharin AYu, Maltsev MV, Palukha UYu, editors. *Computer data analysis and modeling: stochastics and data science. Proceedings of the 13th International conference; 2022 September 6–10; Minsk, Belarus*. Minsk: Belarusian State University; 2022. p. 211–217.
5. Voloshko VA. [On asymptotic properties for a family of χ^2 -tests of pure randomness of binary sequence]. In: Kharin YuS, editor. *Teoreticheskaya i prikladnaya kriptografiya. Materialy II Mezhdunarodnoi nauchnoi konferentsii; 19–20 oktyabrya 2023 g.; Minsk, Belarus* [Theoretical and applied cryptography. Proceedings of the 2nd International conference; 2023 October 19–20; Minsk, Belarus]. Minsk: Belarusian State University; 2023. p. 15–43. Russian.
6. Hayashi M, Watanabe S. Information geometry approach to parameter estimation in Markov chains. *Annals Statistics*. 2016;44(4):1495–1535. DOI: 10.1214/15-AOS1420.
7. National Institute of Standards and Technology. *Security requirements for cryptographic modules*. Gaithersburg: National Institute of Standards and Technology; 2001. 64 p. (Federal information processing standards publications; FIPS PUB 140-2). DOI: 10.6028/NIST.FIPS.140-2.
8. Rukhin A, Soto J, Nechvatal J, Smid M, Barker E, Leigh S, et al. *A statistical test suite for random and pseudorandom number generators for cryptographic applications*. Gaithersburg: National Institute of Standards and Technology; 2010. [131] p. (NIST special publications; SP 800-22 revision 1a).
9. Jordan C. Essai sur la geometrie a n dimensions. *Bulletin de la societe mathematique de France*. 1875;3:103–174.

Received 24.12.2025 / revised 19.01.2026 / accepted 29.01.2026.

ПОСЛЕДОВАТЕЛЬНАЯ СТАТИСТИЧЕСКАЯ ПРОВЕРКА ГИПОТЕЗ ДЛЯ ДИСКРЕТНЫХ СЛУЧАЙНЫХ ДАННЫХ С БЛОЧНОЙ СТРУКТУРОЙ

А. Ю. ХАРИН^{1), 2)}

¹⁾Белорусский государственный университет, пр. Независимости, 4, 220030, г. Минск, Беларусь

²⁾Научно-исследовательский институт прикладных проблем математики и информатики БГУ,
пр. Независимости, 4, 220030, г. Минск, Беларусь

Аннотация. Рассмотрена актуальная математическая задача компьютерного анализа данных – задача последовательной статистической проверки простых гипотез о распределении вероятностей дискретных случайных данных, имеющих блочную структуру. Выведены легко интерпретируемые и удобные для компьютерной реализации явные выражения статистик последовательных тестов (статистических критериев). Разработан подход для вычисления характеристик эффективности решающих правил – вероятностей ошибочных решений и математических ожиданий случайного числа наблюдений, необходимых для обеспечения требуемой точности решений. Получены асимптотические разложения для указанных характеристик эффективности при искажении гипотетической вероятностной модели.

Ключевые слова: случайные дискретные данные; блочная структура; простые гипотезы; последовательный статистический тест; вероятность ошибки; математическое ожидание случайного числа наблюдений; асимптотические разложения.

SEQUENTIAL STATISTICAL HYPOTHESES TESTING FOR DISCRETE RANDOM DATA WITH BLOCK STRUCTURE

A. Yu. KHARIN^{a, b}

^aBelarusian State University, 4 Niezaliezhnasci Avenue, Minsk 220030, Belarus

^bResearch Institute for Applied Problems of Mathematics and Informatics, Belarusian State University,
4 Niezaliezhnasci Avenue, Minsk 220030, Belarus

Abstract. An important mathematical problem of computer data analysis – the problem of sequential statistical testing of simple hypotheses on probability distributions of discrete random data with a block structure – is considered. Explicit expressions of the sequential tests statistics, transparent for interpretation and convenient for computer realisation, are derived. An approach to calculate the performance characteristics – error probabilities of the decisions and mathematical expectations of the random number of observations required to guarantee the requested accuracy for decision rules – is

Образец цитирования:

Харин АЮ. Последовательная статистическая проверка гипотез для дискретных случайных данных с блочной структурой. Журнал Белорусского государственного университета. Математика. Информатика. 2026;1:75–80 (на англ.).
EDN: OPJJSF

For citation:

Kharin AYu. Sequential statistical hypotheses testing for discrete random data with block structure. *Journal of the Belarusian State University. Mathematics and Informatics*. 2026;1: 75–80.
EDN: OPJJSF

Автор:

Алексей Юрьевич Харин – доктор физико-математических наук, профессор; заведующий кафедрой теории вероятностей и математической статистики факультета прикладной математики и информатики¹⁾, главный научный сотрудник научно-исследовательской лаборатории статистического анализа и моделирования²⁾.

Author:

Alexey Yu. Kharin, doctor of science (physics and mathematics), full professor; head of the department of probability theory and mathematical statistics, faculty of applied mathematics and computer science^a, and chief researcher at the laboratory of statistical analysis and modelling^b.
kharinay@bsu.by



developed. Asymptotic expansions for the mentioned performance characteristics under distortion of the hypothetical probability model are constructed.

Keywords: discrete random data; block structure; simple hypotheses; sequential statistical test; error probability; mathematical expectation of the random number of observations; asymptotic expansions.

Introduction

Data is one of the most essential resources of the world economy, and computer data analysis becomes a standard part of the modern life. Efficiency of data analysis defines the success in a growing variety of directions. Discrete data become an extremely important class of data for several reasons. Firstly, discrete data are natural for human being thinking. Secondly, they include binary data that describe many situations in terms of presence – absence, positive – negative, etc. Thirdly, discrete data can be used for description of a significantly rich family of observations and are more clearly processed by computers. Traditional methods of statistical analysis and software are often not applied to such data, as their assumptions are usually not satisfied for the discrete data models, or they are not effective.

As deterministic approach has a limited potential to describe the real-life processes, observed data is usually considered to be random, and probabilistic models are used if mathematics is needed. In these models, an important problem that often appears, is a problem of discrimination between two typical situations on the probability distributions of random data. These two typical situations can be formulated in terms of simple hypotheses on the probability distribution, and the problem turns to the problem of statistical testing of two simple hypotheses [1].

In many fields, especially in statistical quality control, automatic warning systems, personalised medicine and financial decision making, it is important to use the minimal number of observations that guarantee the requested accuracy [2]. Sequential statistical tests [3] are constructed in agreement to this principle with the assumption that the number of observations to be used is defined through the observation process, depend on observations themselves, and thus is a random variable. Due to the complex structure of sequential decision rules, usually analysis of their performance characteristics (error probabilities of decisions and mathematical expectations of the random number of required observations) is moved from theoretical one to the computer simulation [4].

In the paper for the important model of discrete data that has a block structure we develop an approach to calculate and analyse the performance characteristics of sequential tests. In practice the factual probability distribution of data often deviates from the hypothetical one – the hypothetical probability distribution is distorted [5; 6], so we also provide a possibility to calculate the characteristics under distortions of the hypothetical data model.

Sequential probability ratio test for discrete random data with block structure

Denote by $\mathbf{N}$, $\mathbf{Z}$, $\mathbf{R}$ correspondently sets of positive integer, integer and real numbers; $\mathbf{Z}_+ = \mathbf{N} \cup \{0\}$. Let independent identically distributed $K\tau$ -dimensional discrete random vectors with a block structure be observed:

$$X_t = (x_{t,i}) = \begin{pmatrix} X_t^1 \\ X_t^2 \\ \vdots \\ X_t^K \end{pmatrix} \in A^{K\tau}, X_t^k = \begin{pmatrix} x_{t,(k-1)\tau+1} \\ \vdots \\ x_{t,k\tau} \end{pmatrix} \in A^\tau, k = 1, 2, \dots, K, t \in \mathbf{N}, \quad (1)$$

where $A = \{a_1, \dots, a_M\}$ is a finite set with $2 \leq M < +\infty$ elements $a_1 < a_2 < \dots < a_M$; X_t^k is the block number k at the moment t ; $2 \leq K < +\infty$ is a number of blocks; $\tau \in \mathbf{N}$ is a size of the blocks. Let us mention that the equal size for all blocks is set here just to simplify the notation of the results, but for the general case where blocks in the vector X_t may have different sizes the results derivation scheme remains valid.

The blocks $X_t^1, \dots, X_t^K$ are assumed to be independent but the components inside them are stochastically dependent. Because of the block independence in equalities (1), the probability distribution of an observation X_t is separable:

$$P(U) ::= P\{X_t = U\} = \prod_{k=1}^K P^{(k)}(U^k), \quad (2)$$

where $P^k(U^k) ::= P\{X_t^k = U^k\}$ is the probability distribution of the random subvector X_t^k .

Assume that there are two alternatives on the probability distribution (2) in the following parsimonious parametric form:

$$P_v^k(U^k) = P^k(U^k; v) = a^{-J_v^k(U^k)}, U^k \in A^\tau, k = 1, 2, \dots, K, v \in \{0, 1\}, \quad (3)$$

where $a \in \mathbf{R}$, $a > 1$, is some parameter; $J_v^k(U^k) : A^\tau \times \{0, 1\} \rightarrow \mathbf{Z}_+$ is a function that satisfies the normalising condition

$$\sum_{U^k \in A^\tau} a^{-J_v^k(U^k)} = 1. \quad (4)$$

Denote $J_v(U) = \sum_{k=1}^K J_v^k(U^k)$ and $P^v(U) = a^{-J_v(U)}$, $U \in A^{K\tau}$. There are two simple hypotheses considered on the parameters $v \in \{0, 1\}$ of probability distribution (3):

$$H_0: v = 0, H_1: v = 1. \quad (5)$$

Let us denote the following accumulated log-likelihood ratio test statistic for the observed sample of size n :

$$\Lambda_n = \Lambda_n(X_1, \dots, X_n) = \sum_{t=1}^n \lambda_t, \quad (6)$$

where

$$\lambda_t = \lambda(X_t) = \log_a \left(\frac{P_1(X_t)}{P_0(X_t)} \right) = \sum_{k=1}^K \log_a \left(\frac{P_1^k(X_t^k)}{P_0^k(X_t^k)} \right) \quad (7)$$

is the log-likelihood ratio for the observation vector X_t , $t \in \mathbf{N}$.

Denote by $\delta_{i,j}$ the Kronecker delta, by

$$n^k(U^k) = \sum_{t=1}^n \delta_{X_t^k, U^k}, U^k \in A^\tau,$$

the number of the subvectors X_t^k observed within the sample $\{X_1, \dots, X_n\}$ that are equal to U^k ,

$$\sum_{U^k \in A^\tau} n^k(U^k) \equiv n, k = 1, 2, \dots, K.$$

Theorem 1. For the model (1)–(4) the sequence of statistics (6) and (7) for hypotheses (5) is a homogeneous Markov chain with discrete time, and it is represented in the form

$$\Lambda_n = \Lambda_n(X_1, \dots, X_n) = \sum_{k=1}^K \sum_{U^k \in A^\tau} n^k(U^k) (J_0^k(U^k) - J_1^k(U^k)) \in \mathbf{Z}, n \in \mathbf{N}. \quad (8)$$

Proof. The Markov property [7] of the random sequence (6) follows from the independence of its increments (7) due to the independence of random observations (1). Representation (8) is derived by equivalent transformations of sequence of statistics (6) and (7) under assumptions (2)–(4).

Corollary. If under theorem 1 conditions $M = 2$, $K = 1$ ($\tau = 1$, $A = \{0, 1\}$), then test statistic (6) is

$$\Lambda_n = \Lambda_n(x_1, \dots, x_n) = n_0(J_0(0) - J_1(0)) + (n - n_0)(J_0(1) - J_1(1)), n \in \mathbf{N},$$

where n_0 denotes the number of the observations equal to 0.

Using statistics (8), we construct the sequential probability ratio test [3] for hypotheses (5) as follows. The decision after n observations obtained ($n = 1, 2, \dots$) is

$$d = d(n) = \mathbf{1}_{[C_+, +\infty)}(\Lambda_n) + 2 \cdot \mathbf{1}_{(C_-, C_+)}(\Lambda_n), \quad (9)$$

where $\mathbf{1}_D(\cdot)$ denotes the indicator function of a set D . Decisions $d = 0$ and $d = 1$ mean the observation process termination and acceptance of correspondently H_0 or H_1 after n discrete random vectors observed; decision $d = 2$ means that the $(n + 1)$ vector should be observed; $C_-, C_+ \in \mathbf{Z}$, $C_- < C_+$ are parameters of the decision rule (8) called thresholds; in practice they are often calculated according to [3]:

$$C_- = \left\lceil \log_a \left(\frac{\beta_0}{1 - \alpha_0} \right) \right\rceil, C_+ = \left\lceil \log_a \left(\frac{1 - \beta_0}{\alpha_0} \right) \right\rceil,$$

where α_0, β_0 are the admissible values of error type I (to accept H_1 , when H_0 is true) and type II (H_1 is true, H_0 is accepted) probabilities; $[\cdot]$ means the integer part of an argument.

Performance characteristics of the sequential probability ratio test (9)

With thresholds (9) the factual values of error probabilities of type I and II may differ from α_0, β_0 , and the problem of the factual values calculation of the performance characteristics is open for the sequential tests [6; 8–10].

Introduce the notations: $\mathbf{I}_k$ for the identity matrix of size k ; $\mathbf{0}_{m \times n}$ for the zero-matrix of size $(m \times n)$; $\mathbf{1}(\cdot)$ for the unit step function; $\mathbf{1}_k$ for the k -vector column with components equal 1. Denote by $t^{(v)}$ the expected value of the random number of observations (sample size) provided the true hypothesis is $H_v, v \in \{0, 1\}$, and by α, β the fac-

tual values of error type I and II probabilities for test (9). Also, let $N = C_+ - C_-$; $P^{(v)} = \begin{pmatrix} p_{ij}^{(v)} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_2 & \vdots & \mathbf{0}_{2 \times N} \\ \text{---} & \vdots & \text{---} \\ R^{(v)} & \vdots & Q^{(v)} \end{pmatrix}$

be the matrix of size $(N+2) \times (N+2)$ with blocks $R^{(v)}, Q^{(v)}$ defined by

$$p_{ij}^{(v)} = \begin{cases} \sum_{U \in A^{K\tau}} \delta_{J_0(U) - J_1(U), j-i} P^v(U), & i, j \in (C_-, C_+), \\ \sum_{U \in A^{K\tau}} \mathbf{1}(C_- - i + J_1(U) - J_0(U)) P^v(U), & i \in (C_-, C_+), j = C_-, \\ \sum_{U \in A^{K\tau}} \mathbf{1}(J_0(U) - J_1(U) + i - C_+) P^v(U), & i \in (C_-, C_+), j = C_+. \end{cases}$$

Denote

$$\pi^{(v)} = \left(\pi_i^{(v)} \right), \pi_i^{(v)} = \sum_{U \in A^{K\tau}} \delta_{J_0(U) - J_1(U), i} P^v(U), i \in \{C_- + 1, \dots, C_+ - 1\},$$

$$\pi_{C_+}^{(v)} = \sum_{i \geq C_+} \sum_{U \in A^{K\tau}} \delta_{J_0(U) - J_1(U), i} P^v(U),$$

$$\pi_{C_-}^{(v)} = \sum_{i \leq C_-} \sum_{U \in A^{K\tau}} \delta_{J_0(U) - J_1(U), i} P^v(U),$$

$$S^{(v)} = \mathbf{I}_N - Q^{(v)}, B^{(v)} = \left(S^{(v)} \right)^{-1} R^{(v)}, v \in \{0, 1\},$$

and let $W_{(i)}$ means the i column of the matrix W .

Theorem 2. If under the model (1)–(5) $|S^{(v)}| \neq 0, v \in \{0, 1\}$, then the performance characteristics of sequential decision rule (9) are calculated in the explicit form

$$t^{(v)} = \left(\pi^{(v)} \right)' \left(S^{(v)} \right)^{-1} \mathbf{1}_N + 1, \alpha = \left(\pi^{(0)} \right)' B_{(2)}^{(0)} + \pi_{C_+}^{(0)}, \beta = \left(\pi^{(1)} \right)' B_{(1)}^{(1)} + \pi_{C_-}^{(1)}.$$

Proof. Proof is based on the theory of discrete time finite homogeneous Markov chains. The sequence

$$\zeta_n = C_- \cdot \mathbf{1}_{(-\infty, C_-]}(\Lambda_n) + C_+ \cdot \mathbf{1}_{[C_+, +\infty)}(\Lambda_n) + \Lambda_n \cdot \mathbf{1}_{(C_-, C_+)}(\Lambda_n)$$

is a homogeneous Markov chain with $N+2$ states, and C_-, C_+ are the absorbing states.

Case of distorted probability distributions $\{P_v^k(\cdot)\}$

Let $\bar{P}_v(U)$ be the factual probability distribution of the independent observations $X_1, X_2, \dots \in A^{K\tau}$ that may differ from the hypothetical probability distribution $P_v(U)$ determined by probability distribution (3) due to some model distortion.

Denote the sequential statistical test of the form (6) and (9) based on the function

$$\lambda(U) = \log_a \left(\frac{P_1(U)}{P_0(U)} \right), U \in A^{K\tau},$$

by $\delta_\lambda = (\tau_\lambda, d_\lambda)$, where $\tau_\lambda = \inf \{n: d_\lambda(n) \in \{0, 1\}\}$ is the random stopping moment of the test (random number of observations used to accept one of hypotheses (5)). Let for this test

$$\alpha = \alpha(\delta_\lambda) = E_0 \{ \bar{P}_0 \{d_\lambda = 1 | \tau_\lambda\} \}, \beta = \beta(\delta_\lambda) = E_1 \{ \bar{P}_1 \{d_\lambda = 0 | \tau_\lambda\} \} \quad (10)$$

be the factual values of the error type I and type II probabilities, respectively, where $E_v \{ \cdot \}$ means mathematical expectation with the probability distribution $\bar{P}_v$;

$$t^{(v)} = t^{(v)}(\delta_\lambda) = E_v \{ \tau_\lambda \}, v \in \{0, 1\}, \quad (11)$$

is the conditional mathematical expectation of the random size of the sample τ_λ under the true hypothesis is H_k .

Consider the set of performance characteristics $\{\alpha(\delta_\lambda), \beta(\delta_\lambda), t_0(\delta_\lambda), t_1(\delta_\lambda)\}$ for the test δ_λ . Theorem 2 gives expressions for the performance characteristics (10) and (11) of the test δ_λ in case, where the function $\lambda(\cdot)$ defined by the probability distribution (3) and the sequence of statistics (7) has a discrete set of values, and there are no distortions ($\bar{P}_v \equiv P_v$). In the general case for the test δ_λ introduce the notations ($v = 0, 1$)

$$Q^{(v)} = (q_{ij}^{(v)}), q_{ij}^{(v)} = \bar{P}_v \{ \lambda(X_1) = j - i \}, i, j \in \{C_- + 1, \dots, C_+ - 1\},$$

$$R^{(v)} = (r_{ij}^{(v)}), i \in \{C_- + 1, \dots, C_+ - 1\}, j = C_-, C_+,$$

$$r_{iC_-}^{(v)} = \bar{P}_v \{ \lambda(X_1) \leq C_- - i \}, r_{iC_+}^{(v)} = \bar{P}_v \{ \lambda(X_1) \geq C_+ - i \},$$

$$\pi^{(v)} = (\pi_i^{(v)}), \pi_i^{(v)} = \bar{P}_v \{ \lambda(X_1) = i \}, i \in \{C_- + 1, \dots, C_+ - 1\},$$

$$\pi_{C_-}^{(v)} = \bar{P}_v \{ \lambda(X_1) \leq C_- \}, \pi_{C_+}^{(v)} = \bar{P}_v \{ \lambda(X_1) \geq C_+ \}.$$

As it follows from the finite Markov chains theory [7], the condition $|S^{(v)}| \neq 0$ is equivalent to the test finiteness: $\bar{P}_v \{ \tau_\lambda < \infty \} = 1, v \in \{0, 1\}$. This condition is satisfied [2] for case of independent identically distributed random observations, except the case, where $\bar{P}_v \{ \lambda(X_1) = 0 \} = 1$ (in this case $P_0(\cdot) = P_1(\cdot)$).

Let $\underline{\lambda}(U), \bar{\lambda}(U): A^{K\tau} \rightarrow \mathbf{R}$ be some functions, $\delta_{\underline{\lambda}}, \delta_{\bar{\lambda}}$ be the sequential tests based on these functions, $\underline{\Lambda}_n = \sum_{i=1}^n \underline{\lambda}(X_i), \bar{\Lambda}_n = \sum_{i=1}^n \bar{\lambda}(X_i)$. Denote the Markov stopping moments

$$s_0 = \inf \{n: \Lambda_n \leq C_-\}, s_1 = \inf \{n: \Lambda_n \geq C_+\}, \underline{s}_0 = \inf \{n: \underline{\Lambda}_n \leq C_-\},$$

$$\underline{s}_1 = \inf \{n: \underline{\Lambda}_n \geq C_+\}, \bar{s}_0 = \inf \{n: \bar{\Lambda}_n \leq C_-\}, \bar{s}_1 = \inf \{n: \bar{\Lambda}_n \geq C_+\}.$$

Denote also

$$\bar{\alpha} = \alpha(\delta_{\bar{\lambda}}), \bar{\beta} = \beta(\delta_{\bar{\lambda}}), \underline{\alpha} = \alpha(\delta_{\underline{\lambda}}), \underline{\beta} = \beta(\delta_{\underline{\lambda}}),$$

$$Y = \{U \in A^{K\tau}: \lambda(X) \in (C_- - C_+, C_+ - C_-)\}.$$

Let $t^{(v)}$ is used for the mathematical expectation of the random number of observations of the sequential test based on the function $\lambda(\cdot)$ under the true hypothesis H_v , before stopping and accepting one of the hypothesis (5). Also, let $t^{(v)}(\delta_{\bar{\lambda}})$ and $t^{(v)}(\delta_{\underline{\lambda}})$ are used for similar characteristics for the sequential tests based on functions $\bar{\lambda}(\cdot)$ and $\underline{\lambda}(\cdot)$, respectively.

Theorem 3. Let for the sequential statistical test defined above the functions $\underline{\lambda}(\cdot), \bar{\lambda}(\cdot)$ satisfy the conditions

$$\underline{\lambda}(U) \leq \lambda(U) \leq \bar{\lambda}(U) \quad \forall U \in Y,$$

and

$$t^{(v)} < \infty, t^{(v)}(\delta_{\bar{\lambda}}) < \infty, t^{(v)}(\delta_{\underline{\lambda}}) < \infty, v = 0, 1.$$

Then

$$\underline{\alpha} \leq \alpha \leq \bar{\alpha}, \bar{\beta} \leq \beta \leq \underline{\beta}$$

and in the asymptotics $\bar{\alpha} - \underline{\alpha} \rightarrow 0, \underline{\beta} - \bar{\beta} \rightarrow 0$ the following results are valid:

$$t_{-}^{(v)} \leq t^{(v)} \leq t_{+}^{(v)}, v = 0, 1, t_{\pm}^{(v)} = T_{\pm}^{(v)} + o(1),$$

where

$$T_{-}^{(0)} = \bar{\alpha} E_0 \{ \bar{s}_1 | \bar{s}_1 < \bar{s}_0 \} + (1 - \underline{\alpha}) E_0 \{ \underline{s}_0 | \underline{s}_0 < \underline{s}_1 \},$$

$$T_{+}^{(0)} = \underline{\alpha} E_0 \{ \underline{s}_1 | \underline{s}_1 < \underline{s}_0 \} + (1 - \bar{\alpha}) E_0 \{ \bar{s}_0 | \bar{s}_0 < \bar{s}_1 \},$$

$$T_{-}^{(1)} = \underline{\beta} E_1 \{ \underline{s}_0 | \underline{s}_0 < \underline{s}_1 \} + (1 - \bar{\beta}) E_0 \{ \bar{s}_1 | \bar{s}_1 < \bar{s}_0 \},$$

$$T_{+}^{(1)} = \bar{\beta} E_1 \{ \bar{s}_0 | \bar{s}_0 < \bar{s}_1 \} + (1 - \underline{\beta}) E_1 \{ \underline{s}_1 | \underline{s}_1 < \underline{s}_0 \}.$$

Proof. Proof is based on analysis of relations for the introduced Markov moments and for the random events that are generated by them.

Conclusions

An approach to calculate and analyse the performance characteristics of sequential tests for discrete random data is proposed. The block structure in the data makes it possible to consider data not entire observation after entire observation but define the stopping moment on the basis of information available in some blocks only. The results can be applied to construct robust sequential tests [8] for discrete random data and for the model of random sequences with a trend [9]. The results are also potentially applicable to the case of more than two hypotheses [11], complex hypotheses [10] and to the analysis of truncated sequential tests [12].

References

1. Mukhopadhyay N, de Silva B. *Sequential methods and their applications*. New York: Marcel Dekker; 2009. 504 p.
2. Lai T. Sequential analysis: some classical problems and new challenges. *Statistica Sinica*. 2001;11:303–408.
3. Wald A. *Sequential analysis*. New York: John Wiley and Sons; 1947. 212 p.
4. Kharin AY, Tu TT. On error probability calculation for the truncated sequential probability ratio test. *Journal of the Belarusian State University. Mathematics and Informatics*. 2018;1:68–76. Russian.
5. Huber P, Ronchetti E. *Robust statistics*. New York: Wiley; 2009. 384 p.
6. Kharin AY. The influence of distortions of the prior distribution on the characteristics of the sequential test for composite hypotheses. *Journal of Mathematical Sciences*. 2022;267(1):3–10.
7. Kemeny JG, Snell JL. *Finite Markov chains*. New York: Springer; 1960. 224 p.
8. Kharin AY. *Robustness of Bayesian and sequential statistical decision rules*. Minsk: Belarusian State University; 2013. 207 p. Russian.
9. Kharin A, Tu TT. Performance and robustness analysis of sequential hypotheses testing for time series with trend. *Austrian Journal of Statistics*. 2017;46(3–4):23–36. DOI: 10.17713/ajs.v46i3-4.668.
10. Kharin AY. An approach to asymptotic robustness analysis of sequential tests for composite parametric hypotheses. *Journal of Mathematical Sciences*. 2017;227(2):196–203.
11. Tu TT, Kharin AY. Sequential probability ratio test for many simple hypotheses on parameters of time series with trend. *Journal of the Belarusian State University. Mathematics and Informatics*. 2019;1:35–45. Russian. DOI: 10.33581/2520-6508-2019-1-35-45.
12. Kharin AY. Statistical sequential hypotheses testing on parameters of probability distributions of random binary data. *Journal of the Belarusian State University. Mathematics and Informatics*. 2021;2:60–66. DOI: 10.33581/25206508202126066.

Received 24.12.2025 / revised 19.01.2026 / accepted 29.01.2026.

ДИСКРЕТНАЯ МАТЕМАТИКА И МАТЕМАТИЧЕСКАЯ КИБЕРНЕТИКА

DISCRETE MATHEMATICS AND MATHEMATICAL CYBERNETICS

УДК 004.9:656.6

ВКЛЮЧЕНИЕ ДОСТУПНОСТИ БЫСТРЫХ ЗАРЯДНЫХ СТАНЦИЙ В МОДЕЛЬ НАЗНАЧЕНИЙ ГОРОДСКИХ ТРАНСПОРТНЫХ ПОТОКОВ НА ОСНОВЕ АЛГОРИТМА ФРАНКА – ВУЛЬФА

ДУ СИЧЖОУ¹⁾, Д. С. САРАЖИНСКИЙ¹⁾, Д. В. КАПСКИЙ^{1), 2)}

¹⁾Белорусский национальный технический университет,
пр. Независимости, 65, 220013, г. Минск, Беларусь

²⁾Академия управления при Президенте Республики Беларусь,
ул. Московская, 17, 220007, г. Минск, Беларусь

Аннотация. В связи с ростом использования электромобилей возникает необходимость в адаптации транспортных моделей для учета специфики их эксплуатации. Целью работы является модификация классической модели назначений транспортных потоков на основе алгоритма Франк – Вульфа, широко применяемой в транспортном планировании, для корректного отражения процессов зарядки электромобилей в городских условиях. Новизна исследования заключается в разработке подхода, который, в отличие от существующих подходов, комплексно учитывает три ключевые особенности: требование прохождения маршрута электромобиля, нуждающегося в зарядке, ровно через одну зарядную станцию; возможность движения по путям с циклами для доступа к зарядным станциям и последующего следования к пункту назначения в условиях ограничений городской дорожно-транспортной сети; поведенческий аспект отказа пользователя от зарядки, если суммарные временные затраты (включая движение к станции и ожидание в очереди) превышают некоторое пороговое значение. В статье также предлагается функция временных затрат, учитывающая специфику процесса зарядки, в том числе очереди и количество зарядных мест. Результаты демонстрируют, что классическую модель назначений транспортных потоков на основе алгоритма Франк – Вульфа действительно можно модифицировать для учета указанных особенностей. Отмечается,

Образец цитирования:

Ду Сичжоу, Саражинский ДС, Капский ДВ. Включение доступности быстрых зарядных станций в модель назначений городских транспортных потоков на основе алгоритма Франк – Вульфа. *Журнал Белорусского государственного университета. Математика. Информатика.* 2026;1:81–93. EDN: LYDVNK

For citation:

Du Sizhuo, Sarazhinsky DS, Kapski DV. Incorporating fast charging stations accessibility into an urban traffic assignment model based on the Frank – Wolfe algorithm. *Journal of the Belarusian State University. Mathematics and Informatics.* 2026; 1:81–93. Russian. EDN: LYDVNK

Сведения об авторах см. на с. 93.

Information about the authors see p. 93.



Специальный выпуск «Стохастика и анализ данных»
Special issue «Stochastic and Data Science»

что предлагаемая модель может характеризоваться множественностью равновесных распределений потоков, зависящих от начальных условий. Практическая значимость работы состоит в возможности более точного моделирования транспортных потоков и нагрузки на зарядную инфраструктуру в городах.

Ключевые слова: электрокар; быстрая зарядная станция; распределение транспортных потоков по сети; равновесие Вардропа; алгоритм Франк – Вульфа.

INCORPORATING FAST CHARGING STATIONS ACCESSIBILITY INTO AN URBAN TRAFFIC ASSIGNMENT MODEL BASED ON THE FRANK – WOLFE ALGORITHM

DU SIZHUO^a, D. S. SARAZHINSKY^a, D. V. KAPSKI^{a, b}

^aBelarusian National Technical University,
65 Niezaliezhnasci Avenue, Minsk 220013, Belarus

^bAcademy of Public Administration under the President of the Republic of Belarus,
17 Maskowskaja Street, Minsk 220007, Belarus

Corresponding author: D. V. Kapski (d.kapski@gmail.com)

Abstract. The growing use of electric vehicles necessitates the adaptation of transport models to account for the specifics of their operation. The aim of this work is to modify the classic traffic assignment model based on the Frank – Wolfe algorithm, widely used in transportation planning, to accurately reflect electric vehicle charging processes in urban environments. The novelty of the research lies in the development of an approach that, unlike existing approaches, comprehensively addresses three key features: the requirement for an electric vehicle needing to charge to route through exactly one charging station; the possibility of routing along paths with cycles to access charging stations and subsequently reach the destination, considering urban road network constraints; the behavioural aspect of a user's decision to forgo charging if the total time costs (including travel to the station and queuing) exceed a certain threshold. The paper also proposes a time-cost function that accounts for the specifics of the charging process, including queues and the number of charging spots. The results demonstrate that the classic traffic assignment model based on the Frank – Wolfe algorithm can indeed be modified to incorporate these features. It is noted that the proposed model may exhibit multiple equilibrium flow distributions, dependent on initial conditions. The practical significance of the work lies in its potential for more accurate modelling of traffic flows and the load on urban charging infrastructure.

Keywords: electric vehicle; fast charging station; traffic assignment; Wardrop equilibrium; Frank – Wolfe algorithm.

Введение

Устойчивое развитие городов невозможно без создания экологичной и комфортной транспортной системы, имеющей как можно более низкий (желательно нулевой) углеродный след¹. В связи с этим, помимо других мер, активно применяется замена классических (бензиновых) автомобилей на транспортные средства, использующие альтернативные, зеленые виды энергии. Наибольшую популярность получили электромобили в различных вариантах конструктивного исполнения. Рост использования электромобилей как средств обеспечения городской мобильности и специфика данных транспортных средств обуславливают необходимость пересмотра известных транспортных моделей, применяемых в транспортном планировании городов. В частности, требуется включение связанных с эксплуатацией таких автомобилей закономерностей в широко используемую макромодель распределения транспортного спроса по дорожной сети (речь о ней пойдет ниже).

Неформально задачу распределения транспортного спроса по дорожной сети (в том виде, в котором она встречается в транспортном планировании) можно описать следующим образом: пусть имеется дорожная сеть города и известен транспортный спрос, т. е. общее количество поездок, которые жители хотят совершить за определенное время между различными районами города. Для целей транспортного моделирования эти районы представляются дополнительными (специальными) транспортными узлами сети – источниками, откуда поездки начинаются, и стоками, где поездки заканчиваются. Каждый

¹Устойчивая городская мобильность: теория и практика развития : учебник / А. О. Лобашов [и др.]. М. ; Вологда : Инфраинженерия, 2023. 236 с.

источник генерирует заданное число поездок к соответствующему стоку. Когда этот спрос реализуется, он формирует транспортные потоки, которые в модели представляются интенсивностями движения автомобилей по различным маршрутам и сегментам сети (количество автомобилей, проходящих по маршруту или сегменту дорожной сети в единицу времени).

Поскольку от источника к стоку почти всегда существует несколько альтернативных путей, возникает ключевой вопрос: «Каким образом транспортные потоки (водители) распределяются по этим маршрутам?» Это так называемая задача о назначении транспортных потоков в маршрутной (*path-based*) постановке [1]. Поскольку число маршрутов может оказаться очень большим, решение такой задачи бывает вычислительно сложным и часто непрактичным. По этой причине также рассматривают другую задачу: предсказать суммарные транспортные потоки на отдельных сегментах дорожной сети. Это задача о назначении транспортных потоков в сегментной (*link-based*) постановке [1], которая и будет рассматриваться далее.

Теперь более формально представим дорожную сеть в виде простого ориентированного графа (орграфа) (V°, E°) , где V° – вершины, представляющие собой транспортные узлы сети; $E^\circ \subseteq V^\circ \times V^\circ$ – множество дуг, соответствующих односторонним сегментам дорожной сети. (Двустороннее движение по участку дороги моделируется парой противоположно направленных дуг.) Расширим этот орграф до орграфа (V, E) за счет добавления к нему двух непересекающихся множеств вершин – O (множество вершин, отвечающих источникам генерации транспортного спроса) и D (множество вершин, соответствующих стокам), а также множеств дуг $E_O \subseteq O \times V^\circ$ и $E_D \subseteq V^\circ \times D$. (Заметим, что случай, когда один и тот же транспортный узел сети служит как пунктом отправления, так и пунктом назначения поездок, моделируется здесь парой формально различных вершин – одной вершиной из множества O и одной вершиной из множества D .)

Под путем, соединяющим источник $o \in O$ и сток $d \in D$, будем понимать конечную последовательность $(e_1, e_2, \dots, e_n)$ дуг $e_k = (v_k, v'_k) \in E$ такую, что $v_1 = o$, $v'_n = d$ и $v'_{k-1} = v_k$, $k = 2, \dots, n$. Обозначим через ${}^{od}\mathcal{P}$ множество всех таких путей. Положим, что для всякой пары (o, d) , которую также будем называть корреспонденцией, соответствующее множество путей непустое. С учетом вышеопределенного соответствующий орграф договоримся называть орграфом дорожной сети.

С каждой корреспонденцией (o, d) свяжем значение $q_{od} \geq 0$ – транспортный спрос, представляющий собой количество начинающихся в единицу времени поездок из источника $o \in O$ в сток $d \in D$. Эти поездки, как указывалось ранее, порождают физические транспортные потоки по маршрутам и сегментам дорожной сети. В орграфе, моделирующем дорожную сеть, их можно представить через приписывание путям и дугам соответствующих величин интенсивностей движения. В последующем эти величины будем называть транспортными потоками (порожденными транспортным спросом), а процедуру их определения на основе транспортного спроса – распределением спроса и потоков по орграфу дорожной сети. Далее предполагается, что пропускные способности сегментов дорожной сети неограниченны, поэтому от распределения спроса по орграфу дорожной сети будет требоваться только выполнение соотношений, отвечающих сохранению физического транспортного потока.

Для всякого пути $p \in {}^{od}\mathcal{P}$ обозначим через x_p ту часть транспортного спроса, которая распределена на p . Тогда задачу о назначении транспортных потоков в маршрутной постановке можно представить как задачу предсказания значений всех x_p , $p \in {}^{od}\mathcal{P}$, $o \in O$, $d \in D$, при заданном наборе $\mathbf{q} = (q_{od})_{o \in O, d \in D}$.

Чтобы сформулировать задачу о назначении транспортных потоков в сегментной постановке, для всякой дуги $e \in E$ пошагово определим следующие величины:

- транспортный поток по дуге pf_e , порождаемый распределением транспортного спроса на путь $p = (e_1, \dots, e_n)$:

$${}^pf_e = \sum_{k=1}^n x_p \delta_{ee_k}, \quad (1)$$

где $\delta_{ee'} = 1$, если $e' = e$, и $\delta_{ee'} = 0$ в противном случае;

- транспортный поток по дуге ${}^{od}f_e$, порождаемый распределением транспортного спроса корреспонденции (o, d) по путям $p \in {}^{od}\mathcal{P}$:

$${}^{od}f_e = \sum_{p \in {}^{od}\mathcal{P}} {}^pf_e; \quad (2)$$

- общий транспортный поток по дуге f_e :

$$f_e = \sum_{o \in O, d \in D} {}^{od}f_e. \quad (3)$$

С учетом этого задачу о назначении транспортных потоков в сегментной постановке можно сформулировать как задачу предсказания всех $^{od}f_e$ (либо только f_e), $e \in E, o \in O, d \in D$, при заданном q .

Одним из классических подходов к решению указанной задачи является подход, использующий понятие равновесного распределения транспортных потоков по дорожной сети [1–3]: поскольку пользователь вынужден платить за передвижение в первую очередь временем и поскольку типично время передвижения по сегментам дорожной сети зависит от количества пользователей на ней (плотный транспортный поток движется с меньшей скоростью), то можно рассматривать передвижение пользователей в сети как некооперативную игру между ними, где каждый «игрок» стремится «выиграть» себе время за счет выбора подходящего маршрута. Как следствие, естественно ожидать, что если и существует стационарное распределение транспортных потоков по дорожной сети, то оно должно соответствовать известной в теории игр равновесной по Нэшу стратегии. В транспортной области науки это соображение представлено первым принципом Вардропы (принципом равновесия пользователя)² [4]: равновесное распределение транспортных потоков по дорожной сети – это такое распределение, при котором ни одному участнику движения не выгодно в одностороннем порядке изменять свой маршрут. Иными словами, для любой пары источник – сток общая стоимость движения по любому из используемых маршрутов является одинаковой и не превышает стоимости движения по любому неиспользуемому маршруту для этой же пары.

Примечательно, что при довольно естественном во многих ситуациях предположении, согласно которому стоимость передвижения по всякой дуге $e \in E$ зависит только от общего транспортного потока f_e по этой дуге (и не зависит от потоков по другим дугам), задачу поиска равновесного распределения транспортных потоков можно представить в виде задачи оптимизации. Данное обстоятельство обусловлено следующими соображениями [5].

Пусть $c_e = c_e(f)$ – функция стоимости передвижения по дуге $e \in E$ от заданного на ней общего транспортного потока $f \geq 0$, которая монотонно зависит от f , а $C_e(f) = \int_0^f c_e(u) du$ – так называемая интегральная функция стоимости (первообразная функции стоимости).

Распределение транспортных потоков по путям x_p , $p \in {}^{od}\mathcal{P}$, $o \in O, d \in D$, можно представить в виде вектора $\mathbf{x} = (x_p)_{p \in {}^{od}\mathcal{P}, o \in O, d \in D}$ некоторого линейного пространства $\mathcal{L}$. Подмножество этого пространства, каждый элемент которого отвечает некоторому распределению транспортных потоков, порожденных q , договоримся называть допустимым и обозначать через X_q . Согласно формулам (1)–(3) всякий вектор $\mathbf{x} \in X_q$ определяет величины транспортных потоков по дугам ${}^p f_e, {}^{od} f_e$ и $f_e, e \in E, p \in {}^{od}\mathcal{P}, o \in O, d \in D$. Таким образом, можно говорить, что заданы функциональные зависимости между $\mathbf{x}$ и соответствующими потоками. Договоримся эти зависимости обозначать через ${}^p \hat{f}_e = {}^p \hat{f}_e(\mathbf{x})$, ${}^{od} \hat{f}_e = {}^{od} \hat{f}_e(\mathbf{x})$ и $\hat{f}_e = \hat{f}_e(\mathbf{x})$ соответственно.

С учетом вышеизложенного рассмотрим интегральную функцию суммарных потерь на сети

$$Z(\mathbf{x}) = \sum_{e \in E} C_e(\hat{f}_e(\mathbf{x})) = \sum_{e \in E} C_e(f_e), \quad \hat{f}_e = \hat{f}_e(\mathbf{x}). \quad (4)$$

Для удобства проведения анализа далее будем считать, что все введенные выше функции определены своими формулами и для ситуаций $\mathbf{x} \in \mathcal{L} \setminus X_q, f < 0$.

Рассмотрим распределение транспортных потоков $\mathbf{x}^* \in X_q$, доставляющее минимум функции $Z(\mathbf{x})$. Как известно, при всякой вариации (малом изменении) аргумента в точке минимума изменение функции в линейном приближении неотрицательно. Исходя из этого, рассмотрим вариацию $d\mathbf{x} \in \mathcal{L}$, состоящую в том, что часть транспортного потока некоторой корреспонденции (o, d) снимается с одного пути $p^* \in {}^{od}\mathcal{P}$ и перекладывается на другой путь $p \in {}^{od}\mathcal{P}$ (это соответствует ситуации, когда бесконечно малая часть пользователей решили изменить свой путь с p^* на p). С учетом того, что $C'_e(f) = c_e(f)$, можем записать

$$dZ = \sum_{e \in E} c_e(f_e^*) df_e^*, \quad f_e^* = \hat{f}_e(\mathbf{x}^*). \quad (5)$$

Но для рассматриваемого случая $d\hat{f}_e = d({}^{p^*} \hat{f}_e) + d({}^p \hat{f}_e)$, а значит, сумму в выражении (5) можно разбить на две суммы: 1) сумму по дугам пути p^* , содержащую слагаемые $c_e(f_e^*) d({}^{p^*} \hat{f}_e)$; 2) сумму по дугам

²Капский Д. В., Лосин Л. А. Транспорт в планировке городов : учеб.-метод. пособие для студентов специальности 1-44 01 02 «Организация дорожного движения» : в 10 ч. Ч. 1. Транспортное планирование: математическое моделирование. Минск : БНТУ, 2019. 94 с.

пути p , включающую слагаемые $c_e(f_e^*)d({}^p\hat{f}_e)$. Поскольку согласно нашему исходному предположению поток снимается с пути p^* и перекладывается на путь p , то все слагаемые в первой сумме будут отрицательными, а все слагаемые во второй сумме – положительными. Вторую сумму можно трактовать как затраты сменивших путь пользователей на новом пути, а первую сумму – как их затраты при движении по исходному пути. Поскольку суммарное изменение dZ в точке $\mathbf{x}^*$ должно быть неотрицательным, затраты на новом пути будут не меньше затрат на исходном пути, что совпадает с формулировкой принципа равновесия по Вардропу. Таким образом, можем заключить, что каждая точка минимума функции Z соответствует равновесному по Вардропу распределению.

Из приведенных рассуждений и вида выражения (5) вытекает, что если аналогичным образом ввести вектор $\mathbf{f} = ({}^{od}f_e)_{e \in E, o \in O, d \in D}$ с допустимым множеством F_q , рассматривать формулу (3) как определяющую зависимость $\hat{f}_e = \hat{f}_e(\mathbf{f})$ и использовать функцию

$$Z(\mathbf{f}) = \sum_{e \in E} C_e(\hat{f}_e(\mathbf{f})) = \sum_{e \in E} C_e(f_e), f_e = \hat{f}_e(\mathbf{f}), \quad (6)$$

то всякая точка минимума $\mathbf{f}^* \in F_q$ будет соответствовать некоторой точке минимума $\mathbf{x}^* \in X_q$ функции (4). Следовательно, можно говорить о том, что задача поиска равновесного распределения транспортных потоков в сегментной постановке также сводится к задаче оптимизации функции (в данном случае функции (6)).

В довольно естественном для практики предположении, согласно которому Z – выпуклая функция (для этого достаточно, чтобы выпуклостью обладала функция C_e , поскольку сумма выпуклых функций является выпуклой), с учетом того, что множества X_q и F_q являются выпуклыми (и компактными в евклидовой метрике), для поиска минимумов соответствующих функций можно использовать итерационный алгоритм Франк – Вульфа [6–8]. Его итерации включают следующие основные шаги.

Шаг 1. Минимизировать (в пределах допустимого множества) линейную аппроксимацию целевой функции, полученную разложением Тейлора в текущей точке.

Шаг 2. Сдвинуться по направлению к найденной точке на шаг, доставляющий минимум исходной целевой функции на этом направлении.

Примечательно, что в силу вида функции Z первый шаг сводится к простой и эффективной процедуре – поиску для каждой корреспонденции (o, d) пути минимальной стоимости (например, с помощью алгоритма Дейкстры) и временному распределению на него всего спроса q_{od} [1].

Подобная реализация первого шага и сделала этот алгоритм фундаментальным инструментом в транспортном планировании. Именно данный алгоритм поиска равновесного распределения транспортных потоков далее будем относить к (классической) модели назначений транспортных потоков Франк – Вульфа.

При попытке применить указанную модель к сетям с быстрыми зарядными станциями ее ключевые допущения перестают работать, поскольку процесс зарядки вводит новые структурные, маршрутные и поведенческие ограничения, которые не могут быть отражены в рамках классического подхода или его известных вариаций. В частности, возникают следующие особенности: транспортный спрос естественным образом разбивается (по меньшей мере) на два класса – спрос со стороны пользователей, которые нуждаются в зарядке, и спрос со стороны пользователей, не нуждающихся в зарядке. То же происходит и с элементами дорожной сети: среди них начинают выделяться особые элементы (узлы или сегменты), которые ассоциируются с зарядными станциями (типично недоступны для пользователей, не нуждающихся в зарядке). Некоторые эти особенности могут быть напрямую отражены в классической модели назначений транспортных потоков Франк – Вульфа (например, с помощью варианта модели многопользовательского равновесия [5] с двумя классами пользователей, одному из которых запрещено при передвижении использовать выделенные элементы сети, связанные с зарядными станциями), однако ряд особенностей требуют пересмотра модели. Укажем наиболее принципиальные из них.

1. При распределении по орграфу дорожной сети транспортных потоков пользователей, которые решили осуществить зарядку, необходимо обеспечить, чтобы путь у таких потоков включал ровно один элемент сети, отвечающий за зарядку.

2. Следует предусмотреть возможность распределения потоков пользователей, решивших осуществить зарядку, на пути, содержащие циклы (т. е. на пути, в которых одна и та же дуга может встречаться несколько раз). Данное требование связано с тем, что в городских условиях нередко существуют ограничения на движение по улично-дорожной сети к зарядной станции (например, из-за улиц с односторонним движением или перекрестков с запрещенными поворотами), которые вынуждают водителей использовать циклические маршруты.

3. Требуется учесть, что пользователи могут отказываться от зарядки (откладывать ее на потом), если общая потеря времени (включая движение к месту зарядки и ожидание в очереди) превышает некоторое пороговое значение.

Кроме того, хотя это и не столь принципиально, необходимо предложить новую функцию стоимости (потери времени), учитывающую специфику процесса зарядки, а именно очереди на зарядку, количество зарядных мест на станции и разброс времени зарядки.

Среди литературных источников имеется достаточно много работ, посвященных моделированию распределения транспортных потоков с учетом необходимости зарядки (см., например, [9–11]). Однако ни одна из известных авторам публикаций не охватывает сразу все вышеперечисленные особенности с их учетом в модели назначений транспортных потоков Франк – Вульфа. Наиболее близкой по рассматриваемой проблематике к настоящей работе является статья [12], однако в ней, во-первых, не учитываются требования 1–3, а во-вторых, используется довольно грубая ценовая функция, включающая технические (не имеющие содержательной интерпретации) параметры. Кроме того, модель представлена в виде общей оптимизационной задачи выпуклого программирования с ограничениями, что ставит под сомнение возможность ее масштабирования на сети большого размера (в отличие от модели назначений транспортных потоков Франк – Вульфа, у которой с этим нет принципиальных проблем).

Все вышеизложенное и делает актуальным вопрос о модификации (совершенствовании) модели назначений транспортных потоков Франк – Вульфа. Целью данной работы является разработка соответствующей модификации.

Разработка модифицированной модели назначений транспортных потоков Франк – Вульфа

В этом разделе мы, опираясь на соображения, на которых строится классическая модель назначений транспортных потоков Франк – Вульфа, проведем ряд обоснованных модификаций, чтобы отразить в них требования 1–3 из предыдущего раздела.

Пусть имеется классическая модель назначений транспортных потоков Франк – Вульфа, которая включает:

- модель транспортных потоков, заданную посредством орграфа дорожной сети (V, E) , набора монотонно возрастающих дифференцируемых функций стоимости передвижения $c_e = c_e(f)$, $e \in E$, и набора величин транспортного спроса по корреспонденциям $q = (q_{od})_{o \in O, d \in D}$;

- алгоритм Франк – Вульфа для поиска равновесного распределения транспортных потоков в указанной модели транспортных потоков.

Будем модифицировать модель, чтобы учесть определенную выше специфику, связанную с зарядными станциями.

Модификация модели транспортных потоков. Пусть имеется некоторое множество S , элементы которого представляют зарядные станции. Определим вначале операцию инсталлирования зарядной станции $s \in S$ на дугу $e = (v', v'') \in E$ как операцию, предполагающую следующие последовательные действия:

- 1) $V \mapsto V \cup \{s\}$ (добавить вершину, соответствующую зарядной станции s);
- 2) $E \mapsto E \cup \{(v', s), (s, v'')\}$ (добавить дуги, ведущие к зарядной станции s и от нее);
- 3) приписать добавленным дугам функции стоимости в соответствии со следующими соображениями.

Дуга $a^\uparrow = (v', s)$, ведущая к зарядной станции, моделирует подъезд к самой станции (после съезда с общей дороги), поэтому предусматривает движение только для автомобилей, которые нуждаются в зарядке. Поскольку интенсивности таких потоков предполагаются не слишком большими, то можно не учитывать их влияние на скорость движения по дуге. Соответственно, можно рассматривать эту часть потери времени как постоянную, равную некоторой константе t_a^o (время, необходимое для того, чтобы доехать до станции от узла v'). Далее к этому времени нужно добавить время, связанное с обслуживанием на станции. Для его расчета можно использовать результаты теории массового обслуживания для системы типа $M/G/k$, а именно будем отталкиваться от аналога формулы Поллачека – Хинчина для среднего времени ожидания в очереди $W_q^{M/G/k}$, например, в приближении, описанном в книге [13]:

$$W_q^{M/G/k} \approx W_q^{M/M/k} \frac{1 + J^2}{2} = \eta \tau \frac{1 + J^2}{2},$$

$$\eta = \frac{(k\rho)^k}{k(1-\rho)^2 k! \left(\frac{(k\rho)^k}{(1-\rho)k!} + \sum_{i=0}^{k-1} \frac{(k\rho)^i}{i!} \right)}, \quad \rho = \frac{f\tau}{k} < 1,$$

где τ – среднее время обслуживания одного канала (среднее время зарядки электромобиля); $J = \frac{\sigma}{\tau}$ – коэффициент вариации времени обслуживания (времени зарядки) при стандартном отклонении σ ; f – интенсивность потока прибытия заявок на обслуживание (интенсивность прибытия электромобилей на зарядку); k – количество параллельных обслуживаний (количество зарядных мест на станции).

Поскольку $W_q^{M/M/k}$ как функция от f является строго возрастающей (и строго выпуклой, как следует из публикации [14]) функцией, то это же будет справедливым и для $W_q^{M/G/k}$, что делает последнюю функцию хорошим кандидатом на использование в качестве функции стоимости (поскольку эти свойства гарантирует выпуклость интегральной функции стоимости). Однако она не определена для $\rho \geq 1$, что может приводить к нежелательным последствиям в итерационных алгоритмах наподобие алгоритма Франк – Вульфа. По этой причине предлагается выполнить для нее процедуру регуляризации следующего вида.

Пусть имеется определенная на полуинтервале $[0, 1)$ неограниченная дифференцируемая функция $g = g(\rho)$. Соответствующей ей линейно регуляризованной после точки $\rho_0 \in (0, 1)$ функцией будем называть определенную на полуинтервале $[0, \infty)$ функцию вида

$$[g]_{\rho_0}(\rho) = \begin{cases} g(\rho), & \rho < \rho_0, \\ g(\rho_0) + g'(\rho_0)(\rho - \rho_0), & \rho \geq \rho_0. \end{cases}$$

Очевидно, что такая регуляризация сохраняет строгую монотонность (и выпуклость, хотя уже и не строго), а также дифференцируемость функции.

С учетом сказанного, выбрав достаточно малое $\varepsilon > 0$, окончательно можно записать следующее выражение для функции стоимости:

$$c_{a^\uparrow}(f) = t_{a^\uparrow}^o + [W_q^{M/G/k}]_{1-\varepsilon}(f) + \tau_s,$$

в которой параметры τ_s и σ_s отвечают среднему и стандартному отклонению времени зарядки на станции s соответственно.

Для дуги $a^\downarrow = (s, v'')$, ведущей от станции, имеет смысл учесть только не зависящее от интенсивности потока время движения $t_{a^\downarrow}^o$ и, следовательно, положить, что $c_{a^\downarrow}(f) \equiv t_{a^\downarrow}^o$ (либо внести эту константу в величину $t_{a^\uparrow}^o$, после положив, что $c_{a^\downarrow}(f) \equiv 0$).

С учетом добавления указанной операции введем процедуру модификации модели транспортных потоков как последовательность следующих действий:

- 1) установить моделируемые станции на назначенных им транспортных дугах из множества E^o ;
- 2) дополнительно установить фиктивные станции на всех дугах, ведущих к стокам, т. е. на дугах из множества E_D (рис. 1). Эти станции вводятся для того, чтобы смоделировать наличие других, находящихся за пределами рассматриваемой дорожной сети зарядных станций, пользоваться которыми могут предпочесть водители, если дополнительная потеря времени (на то, чтобы добраться до обычной станции и начать процесс зарядки) превысит некоторое пороговое значение T^{tol} . По этой причине для ценовых функций этих дуг предполагается установка таких параметров, которые в итоге позволят получить равенства

$$c_{a^\uparrow}(f) = T^{tol} + \tilde{\tau}_s, \quad c_{a^\downarrow}(f) \equiv 0, \quad (7)$$

где $\tilde{\tau}_s$ – некоторое типичное для зарядных станций среднее время зарядки (отражает оценку потери времени пользователями на зарядку на типичной станции);

- 3) начать рассматривать набор величин транспортного спроса по корреспонденциям $\mathbf{q} = (q_{od})_{o \in O, d \in D}$ как полученный в результате (поэлементного) суммирования наборов ${}^e\mathbf{q} = ({}^eq_{od})_{o \in O, d \in D}$ и ${}^c\mathbf{q} = ({}^cq_{od})_{o \in O, d \in D}$, которые соответствуют спросу на передвижение пользователей с потребностью в зарядке (тип пользователей e) и пользователей без такой потребности (тип пользователей c), иными словами, начать предполагать выполнение равенства $\mathbf{q} = {}^e\mathbf{q} + {}^c\mathbf{q}$;

4) для каждой корреспонденции (o, d) во множестве всех путей ${}^{od}\mathcal{P}$ выделить множество ${}^e({}^{od}\mathcal{P})$, включающее только пути, проходящие ровно через одну зарядную станцию (пути пользователей типа e), и множество ${}^c({}^{od}\mathcal{P})$, включающее только пути, не проходящие ни через одну зарядную станцию (пути пользователей типа c) (отметим, что эти два множества не пересекаются). Положить, что потоки пользователей с потребностью в зарядке могут двигаться только по путям из множества ${}^e({}^{od}\mathcal{P})$, а потоки пользователей без потребности в зарядке – только по путям из множества ${}^c({}^{od}\mathcal{P})$.

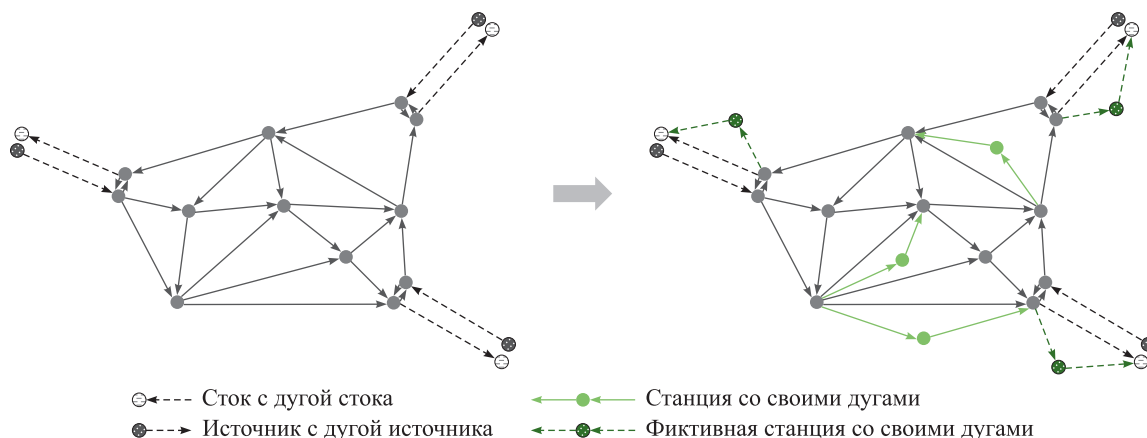


Рис. 1. Схема модификации орграфа дорожной сети

Fig. 1. Road network digraph modification scheme

Модификация алгоритма Франк – Вульфа для поиска равновесного распределения транспортных потоков. С учетом выполненной модификации модели транспортных потоков далее в линейном пространстве $\mathcal{L}$ выделим подмножество $X_{e,q}$ (подмножество X_c), каждый элемент которого $\mathbf{x} = (x_p)_{p \in {}^{od}\mathcal{P}, o \in O, d \in D}$ отвечает некоторому распределению транспортных потоков, порожденных ${}^e q$ (порожденных ${}^c q$), причем $x_p = 0$, если путь p не входит во множество ${}^e({}^{od}\mathcal{P})$ (не входит во множество ${}^c({}^{od}\mathcal{P})$). С учетом этого рассмотрим множество $X_{e,c,q} = X_{e,q} + X_{c,q}$ (где понимается поэлементное сложение). Нетрудно видеть, что $X_{e,c,q}$ – выпуклое (и компактное в евклидовой метрике) множество, и для него $X_{e,c,q} \subseteq X_q$, $\mathbf{q} = {}^e \mathbf{q} + {}^c \mathbf{q}$. Соответственно, для $X_{e,c,q}$ можно провести те же рассуждения, которые проводились выше для X_q , в частности прийти к выводу, что всякая точка $\mathbf{x}^* \in X_{e,c,q}$, доставляющая минимум функции $Z(\mathbf{x})$, определяет равновесное распределение транспортных потоков в модифицированной модели транспортных потоков, причем для его поиска может использоваться алгоритм Франк – Вульфа. Аналогично, как и ранее, для векторов $\mathbf{f} = ({}^{od}f_e)_{e \in E, o \in O, d \in D}$ можно ввести на основе $X_{e,c,q}$ допустимое множество $F_{e,c,q}$ и сделать те же выводы относительно поиска равновесных распределений. Все вышеизложенное обосновывает возможность использования классического подхода (через алгоритм Франк – Вульфа) к поиску равновесных распределений в ситуации, учитывающей специфику процессов, связанных с зарядными станциями, с отдельными оговорками, про которые будет сказано ниже.

Обозначим через M модифицированную модель транспортных потоков и введем следующие функции.

Функция поиска пути минимальной стоимости (*shortest path constrained, SPC*)

Вход: модифицированная модель транспортных потоков M , распределение транспортных потоков $\mathbf{f} \in F_{e,c,q}$, корреспонденция (o, d) , $o \in O$, $d \in D$, тип пути $k \in \{e, c\}$.

Выход: кратчайший по суммарной стоимости путь p заданного типа из o в d при заданных распределениях транспортных потоков в модели.

Как и в исходном алгоритме Франк – Вульфа, эта функция может быть реализована на основе эффективных алгоритмов поиска кратчайшего пути во взвешенном орграфе, в частности, с помощью алгоритма Дейкстры в модифицированной версии, когда дополнительно накладываются ограничения на тип пути, который нужно найти. Сделать это можно, например, заменяя в классическом алгоритме просмотр исходных вершин v орграфа просмотром состояний (v, k_e) , где k_e – количество вершин зарядных станций,

встреченных на пути минимальной стоимости до вершины v . Данное обстоятельство позволяет не только учесть ограничения, но и реализовать построение пути с контурами, когда того требует ситуация.

Функция построения начального распределения транспортных потоков (*initial flow distribution, IFD*)

Вход: модифицированная модель транспортных потоков M .

Выход: распределение транспортных потоков $f \in F_{e,c,q}$.

Выбор реализации данной функции в классической модели достаточно произволен. К наиболее простым вариантам можно отнести: 1) вариант с использованием распределения транспортных потоков по принципу «все или ничего» (*all-or-nothing assignment*), когда весь спрос распределяется на единственный путь, например путь минимальной стоимости; 2) вариант с инкрементальным распределением (*incremental assignment*), когда исходный спрос разбивается на малые части, для каждой из которых производится распределение на путь минимальной стоимости, после чего выполняется перерасчет стоимостей. Произвольность выбора обычно связана с тем фактом, что если целевая функция строго выпуклая (как это обычно предполагается в классической модели), то точка минимума является единственной. Однако в предлагаемой модификации из-за наличия не зависящих от величин транспортных потоков функций стоимостей (7) целевая функция Z уже не будет строго выпуклой (хотя и продолжит быть выпуклой). По этой причине множество точек $Opt_D(Z)$ допустимого множества D , доставляющих минимум целевой функции Z , уже не должно состоять только из одной точки. Более того, итерационная процедура алгоритма не обязана давать последовательность, сходящуюся только к одной точке (а может давать последовательность, имеющую в качестве предельных сразу несколько точек из множества $Opt_D(Z)$ [7]). В связи с этим выбор начальной точки становится важным. Соответственно, выбор варианта реализации функции IFD также важен. По мнению авторов, в качестве наиболее простого и в то же время адекватного варианта можно рассматривать вариант с инкрементальным распределением, поскольку он во многом отражает реальную ситуацию с постепенным ростом спроса в течение суток, начиная с утренних часов. Однако в дальнейшем эта ситуация заслуживает отдельного исследования.

С учетом введенных функций можем описать модифицированный алгоритм Франк – Вульфа следующим образом.

Алгоритм (модифицированный) Франк – Вульфа

Вход: модифицированная модель транспортных потоков M , величина допуска относительной сходимости³ $\epsilon_r > 0$.

Выход: (приближенное⁴) равновесное распределение транспортных потоков $f \in F_{e,c,q}$.

1. Задать начальное распределение $f^{(0)} \in F_{e,c,q}$:

$$f^{(0)} \leftarrow \text{IFD}(M).$$

2. Для каждого $i = 1, 2, \dots$ выполнить следующие действия.

2.1. Определить распределение $y \in F_{e,c,q}$, задающее направление поиска, с помощью процедуры распределения транспортных потоков по принципу «все или ничего».

2.1.1. Положить, что $y \leftarrow 0$.

2.1.2. Для каждого $o \in O, d \in D, \kappa \in \{e, c\}$:

а) найти путь минимальной стоимости p :

$$p \leftarrow \text{SPC}(M, f^{(i-1)}, (o, d), \kappa);$$

б) пересчитать y в предположении, что на найденный путь $p = (e_1, \dots, e_n)$ дополнительно распределится весь спрос q_{od} данного типа пользователей κ , а именно положить, что

$${}^{od}y_e \leftarrow {}^{od}y_e + \sum_{k=1}^n q_{od} \delta_{ee_k}, \quad e \in E$$

(здесь допускается путь с циклами).

³Под этим понятием здесь подразумевается пороговое значение относительной ошибки между последовательными приближениями искомого решения, при достижении которого и ниже которого итерационный процесс считается сошедшимся.

⁴Алгоритм является итерационным, поэтому возвращает приближенное значение равновесного распределения транспортных потоков.

- 2.2. Определить величину шага $\alpha^{(i)} \in [0, 1]$ в направлении поиска путем решения вспомогательной задачи одномерной оптимизации:

$$\alpha^{(i)} = \arg \min_{\alpha \in [0, 1]} Z(f(\alpha)), \quad f(\alpha) = f^{(i-1)} + \alpha(y - f^{(i-1)}).$$

- 2.3. Обновить потоки:

$$f^{(i)} \leftarrow f^{(i-1)} + \alpha^{(i)}(y - f^{(i-1)}).$$

- 2.4. Если относительная разность $Z(f^{(i)})$ и $Z(f^{(i-1)})$ по абсолютной величине меньше ϵ_r , то завершить алгоритм с выводом в качестве результата $f^{(i)}$.

Таким образом, получена модифицированная модель назначений транспортных потоков Франк – Вульфа. Следует отметить, что в применяемых для городов моделях назначений транспортных потоков часто используется подход, который, помимо стоимости движения по дуге, задает стоимость перехода («поворота») с одной дуги на другую [5]. Однако поскольку каждый такой орграф с «поворотами» можно свести к обычному орграфу по схеме, представленной на рис. 2, то можно говорить, что все соображения относительно построения модификации модели остаются справедливыми и в этом случае.

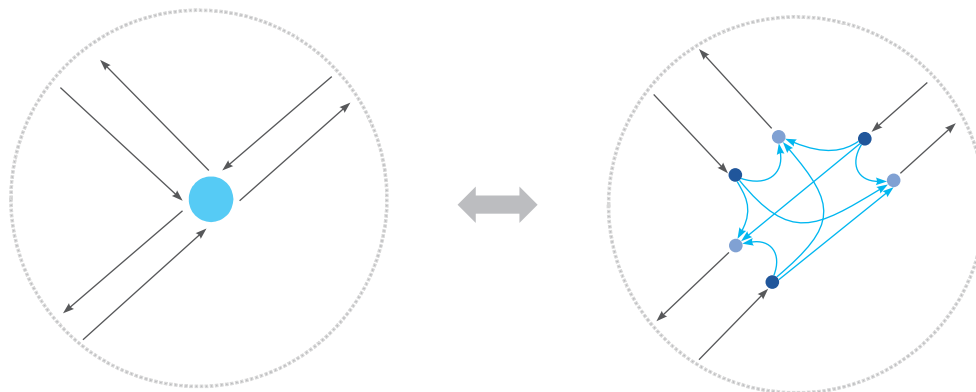


Рис. 2. Схема сведения орграфа с «поворотами» к обычному орграфу
Fig. 2. Scheme of reducing a digraph with «turns» to a standard digraph

Иллюстрация применения модифицированной модели назначений транспортных потоков Франк – Вульфа

В качестве наглядного примера применения предложенной модели рассмотрим задачу оценки загрузки зарядных станций для заданного варианта (конфигурации) их размещения на дорожной сети, а именно модель транспортных потоков с орграфом дорожной сети вида, как на рис. 3 (прообраз фрагмента улично-дорожной сети центральной части г. Минска), в варианте с «поворотами». Во избежание излишней сложности изложения рассматриваем упрощенный случай, когда все повороты разрешены, кроме разворотов; потерями времени на поворотах пренебрегаем. Для дуг, исходящих из источников и ведущих в стоки, устанавливается нулевая длина. Для всех транспортных дуг в качестве функции стоимости используется так называемая функция BPR (*Bureau of Public Roads*, США) [15]:

$$c_e(f) = \frac{l_e}{v_e^0} \left(1 + \alpha \left(\frac{f}{f_e^{\max}} \right)^\beta \right),$$

где l_e – длина дуги; v_e^0 – скорость свободного движения по дуге (в нашем примере 50 км/ч); $f_e^{\max}$ – пропускная способность дорожного сегмента (в нашем примере соответствует величине 1800 автомобилей в час на полосу движения); α , β – калибровочные параметры (в нашем примере выбраны типичные согласно источнику [15] значения $\alpha = 0,15$, $\beta = 4$).

Все станции имеют следующие параметры: $\tau_s = 1800$ с, $\sigma_s = 600$ с, $k_s = 8$. Для всех фиктивных станций $T^{\text{tol}} = 600$ с, $\tilde{\tau}_s = 1800$ с. Для подводящих дуг станций $t_{a \uparrow}^o = 0$.

Предполагаем, что доля пользователей, которые нуждаются в зарядке, является одинаковой для всех корреспонденций и составляет 1 % от общего транспортного спроса. Результирующий транспортный спрос нуждающихся в зарядке ${}^e q = ({}^e q_{od})_{o \in O, d \in D}$ (с совокупной величиной 148 автомобилей в час) графически представлен на рис. 4 (величины общего спроса были получены на основе измерений для упоминавшейся выше части г. Минска в дневное время), где ширина полос пропорциональна величине спроса для каждой корреспонденции.

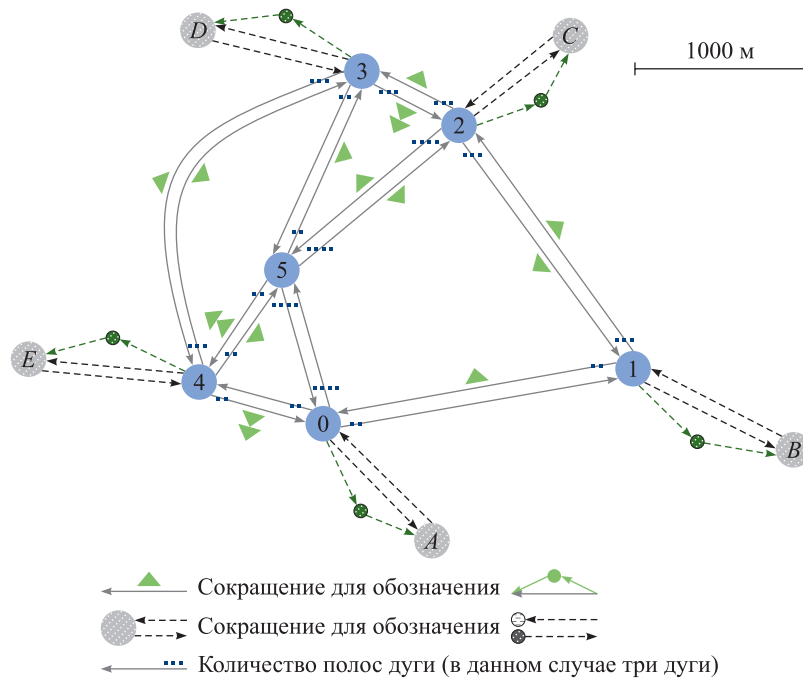


Рис. 3. Тестовый оргграф дорожной сети
Fig. 3. Test road network digraph

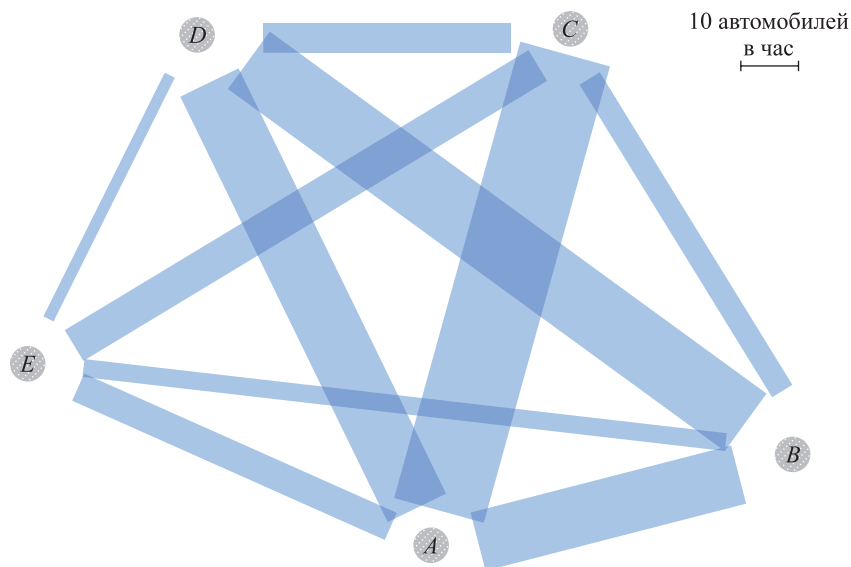


Рис. 4. Картограмма спроса на зарядку
Fig. 4. Charging demand flow map

При расчете использовался вариант IFD с инкрементальным распределением с количеством частей 10. Допуск относительной сходимости составляет $\epsilon_r = 0,1$.

Расчет завершился после 15 итераций с относительным изменением (по последовательным двум итерациям) значений целевой функции 0,000 08 и с максимальным по всем дугам абсолютным изменением (по последовательным двум итерациям) величин потоков $5 \cdot 10^{-5}$ автомобилей в секунду, или 0,15 автомобиля в час.

Особый интерес представляют следующие полученные результаты:

- доля электромобилей, нуждающихся в зарядке, которые не воспользовались ни одной из станций, составляет $\frac{37}{148} \approx 0,25$ (25 %);

- распределение уровня загруженности (коэффициента использования) станций по квартилям $Q1$ (25 %), $Q2$ (50 %), $Q3$ (75 %) имеет вид 0,870, 0,875, 0,880 соответственно.

Исходя из этого, можно сделать следующие выводы. Видно, что заданная конфигурация размещения станций обеспечивает достаточно равномерную нагрузку на них (можно рассматривать в качестве положительного момента). Однако эту конфигурацию трудно признать оптимальной по следующим причинам: она включает 16 зарядных станций с 8 зарядными местами, каждое из которых способно (в пределах) обслуживать 2 автомобиля в час. Теоретическая предельная пропускная способность такой сети составляет 256 автомобилей в час, что значительно больше совокупного спроса на зарядку (148 автомобилей в час). Следовательно, теоретически все нуждающиеся в зарядке электромобили могли быть обслужены, но поскольку этого не произошло (имеем 25 % необслуженных электромобилей), можно говорить о том, что потенциал зарядных станций не был полностью использован, поэтому имеет смысл искать другие варианты размещения станций.

Заключение

В настоящей работе модифицирована классическая модель назначений транспортных потоков на основе алгоритма Франк – Вульфа. Предложенная модификация позволяет учесть ключевые особенности, связанные со спецификой использования электромобилей и процесса их зарядки в условиях городской дорожной сети. В частности, модель интегрирует три принципиальных аспекта: ограничение на прохождение маршрута ровно через одну зарядную станцию, возможность построения путей с циклами для доступа к станциям при наличии транспортных ограничений, а также поведенческий аспект отказа пользователя от зарядки при превышении суммарными временными затратами порогового значения.

Практическая значимость предложенной модели была продемонстрирована в ходе вычислительного эксперимента. Результаты показали, что даже при теоретически достаточной совокупной пропускной способности зарядной инфраструктуры особенности равновесного распределения потоков и поведенческие факторы могут приводить к тому, что значительная доля электромобилей (в рассмотренном примере 25 %) не обслуживаются. Этот вывод подчеркивает важность применения моделей, способных отражать подобные нелинейные эффекты, для адекватной оценки нагрузки на инфраструктуру.

Проведенное исследование открывает перспективы для дальнейших разработок. Предложенная в данной статье модификация модели назначений транспортных потоков может служить ядром для создания алгоритма поиска оптимального расположения зарядных станций для электромобилей в городах, в том числе мегаполисах. Такой алгоритм позволит не просто оценивать заданные конфигурации, а целенаправленно находить решения, минимизирующие как временные затраты водителей, так и долю необслуженного спроса на зарядку.

Библиографические ссылки

1. Fukushima M. A modified Frank – Wolfe algorithm for solving the traffic assignment problem. *Transportation Research. Part B, Methodological*. 1984;18(2):169–177. DOI: 10.1016/0191-2615(84)90029-8.
2. Dafermos SC, Sparrow FT. The traffic assignment problem for a general network. *Journal of Research of the National Bureau of Standards. Section B, Mathematical Sciences*. 1969;73B(2):91–118. DOI: 10.6028/JRES.073B.010.
3. Li F, Wang S. Determining route traffic flows for traffic assignment problem with Frank – Wolfe algorithm. In: Institute of Electrical and Electronics Engineers. *Proceedings of the 2005 IEEE intelligent vehicles symposium; 2005 June 6–8; Las Vegas, USA*. [S. l.]: Institute of Electrical and Electronics Engineers; 2005. p. 669–673. DOI: 10.1109/IVS.2005.1505180.
4. Wardrop JG. Some theoretical aspects of road traffic research. *Proceedings of the Institution of Civil Engineers*. 1952;1(3): 325–362. DOI: 10.1680/ipeds.1952.11259.
5. Shvetsov VI. Mathematical modeling of traffic flows. *Automation and Remote Control*. 2003;64(11):1651–1689. DOI: 10.1023/A:1027348026919.
6. Pokutta S. The Frank – Wolfe algorithm: a short introduction. *Jahresbericht der Deutschen Mathematiker-Vereinigung*. 2024; 126(1):3–35. DOI: 10.1365/s13291-023-00275-x.
7. Bolte J, Combettes CW, Pauwels E. *The iterates of the Frank – Wolfe algorithm may not converge* [Internet]. Toulouse: Toulouse School of Economics; 2022 [cited 2025 May 7]. 15 p. (TSE working papers; number 1311). Available from: https://www.tse-fr.eu/sites/default/files/TSE/documents/doc/wp/2022/wp_tse_1311.pdf.
8. Frank M, Wolfe P. An algorithm for quadratic programming. *Naval Research Logistics Quarterly*. 1956;3(1–2):95–110. DOI: 10.1002/nav.3800030109.

9. Froger A, Jabali O, Mendoza JE, Laporte G. The electric vehicle routing problem with capacitated charging stations. *Transportation Science*. 2022;56(2):460–482. DOI: 10.1287/trsc.2021.1111.
10. Graf L, Harks T, Palkar P. Dynamic traffic assignment for electric vehicles. *Transportation Research. Part B, Methodological*. 2025;195:103207. DOI: 10.1016/j.trb.2025.103207.
11. Liu B, Pantelidis TP, Tam S, Chow JYJ. An electric vehicle charging station access equilibrium model with $M/D/C$ queueing. *International Journal of Sustainable Transportation*. 2023;17(3):228–244. DOI: 10.1080/15568318.2022.2029633.
12. Ferro G, Minciardi R, Parodi L, Robba M. Optimal planning of charging stations and electric vehicles traffic assignment: a bi-level approach. *IFAC – Papers Online*. 2020;53(2):13275–13280. DOI: 10.1016/j.ifacol.2020.12.157.
13. Gross D, Shortle JF, Thompson JM, Harris CM. *Fundamentals of queueing theory*. 4th edition. Hoboken: John Wiley & Sons; 2008. XIII, 509 p. (Wiley series in probability and statistics). DOI: 10.1002/9781118625651.
14. Grassmann W. The convexity of the mean queue size of the $M/M/c$ queue with respect to the traffic intensity. *Journal of Applied Probability*. 1983;20(4):916–919. DOI: 10.2307/3213605.
15. Sheffi Y. *Urban transportation networks: equilibrium analysis with mathematical programming methods*. Englewood Cliffs: Prentice-Hall; 1985. XVI, 399 p.

Получена 30.05.2025 / исправлена 24.11.2025 / принята 24.02.2026.
Received 30.05.2025 / revised 24.11.2025 / accepted 24.02.2026.

Авторы:

Ду Сичжоу – аспирант кафедры транспортных систем и технологий автотракторного факультета. Научный руководитель – Д. В. Капский.

Денис Сергеевич Саражинский – кандидат физико-математических наук, доцент; доцент кафедры транспортных систем и технологий автотракторного факультета.

Денис Васильевич Капский – доктор технических наук, профессор; профессор кафедры транспортных систем и технологий автотракторного факультета¹⁾, профессор кафедры управления информационными ресурсами Института управленческих кадров²⁾.

Authors:

Du Sizhuo, postgraduate student at the department of transport systems and technologies, automotive and tractor faculty. duxinzhou8@gmail.com

<https://orcid.org/0000-0003-4285-0591>

Denis S. Sarazhinsky, PhD (physics and mathematics), docent; associate professor at the department of transport systems and technologies, automotive and tractor faculty. sarazhinsky@mail.ru

Denis V. Kapski, doctor of science (engineering), full professor; professor at the department of transport systems and technologies, automotive and tractor faculty^a, and professor at the department of information resources management, Institute of Managerial Personnel^b.

d.kapsky@gmail.com

<https://orcid.org/0000-0001-9300-3857>

О ХОРОШО ve -УКРЫТЫХ И ХОРОШО ev -УКРЫТЫХ ГРАФАХ

Ю. Л. ОРЛОВИЧ¹⁾, Н. А. ШУТРО¹⁾

¹⁾Белорусский государственный университет, пр. Независимости, 4, 220030, г. Минск, Беларусь

Аннотация. Исследуются классы хорошо ve -укрытых графов и хорошо ev -укрытых графов. Граф называется хорошо ve -укрытым (соответственно хорошо ev -укрытым), если все его независимые минимальные ve -доминирующие (соответственно независимые минимальные ev -доминирующие) множества имеют одинаковую мощность. Показано, что задачи распознавания рассматриваемых классов графов являются со-NP-полными даже для некоторых сужений этих классов. Найдены характеристики в терминах запрещенных порожденных подграфов для максимальных наследственных подклассов хорошо ve -укрытых графов и хорошо ev -укрытых графов. Установлена вычислительная сложность задач, связанных с наименьшими независимыми ve -доминирующими и наименьшими независимыми ev -доминирующими множествами.

Ключевые слова: граф; независимое ve -доминирующее множество; независимое ev -доминирующее множество; хорошо ve -укрытый граф; хорошо ev -укрытый граф; наследственный класс графов; NP-полнота.

ON WELL ve -COVERED AND WELL ev -COVERED GRAPHS

Y. L. ORLOVICH^a, M. A. SHUTRO^a

^aBelarusian State University, 4 Niezaliezhnasci Avenue, Minsk 220030, Belarus

Corresponding author: Y. L. Orlovich (orlovich@bsu.by)

Abstract. In this paper, the classes of well ve -covered graphs and well ev -covered graphs are investigated. A graph is called well ve -covered (well ev -covered, respectively) if all its minimal independent ve -dominating (minimal independent ev -dominating, respectively) sets have the same cardinality. It is shown that the problems of recognising the graph classes under consideration are co-NP-complete even for some restrictions of these classes. Characterisations, in terms of forbidden induced subgraphs, for the maximal hereditary subclasses of well ve -covered graphs and well ev -covered graphs are found. The computational complexity of problems related to minimum independent ve -dominating and minimum independent ev -dominating sets is established.

Keywords: graph; independent ve -dominating set; independent ev -dominating set; well ve -covered graph; well ev -covered graph; hereditary graph class; NP-completeness.

Образец цитирования:

Орлович ЮЛ, Шутро НА. О хорошо ve -укрытых и хорошо ev -укрытых графах. Журнал Белорусского государственного университета. Математика. Информатика. 2026;1: 94–107 (на англ.).
EDN: JOVAZV

For citation:

Orlovich YL, Shutro MA. On well ve -covered and well ev -covered graphs. Journal of the Belarusian State University. Mathematics and Informatics. 2026;1:94–107.
EDN: JOVAZV

Авторы:

Юрий Леонидович Орлович – кандидат физико-математических наук, доцент; декан факультета прикладной математики и информатики.
Никита Алексеевич Шутро – студент факультета прикладной математики и информатики. Научный руководитель – Ю. Л. Орлович.

Authors:

Yury L. Orlovich, PhD (physics and mathematics), docent; dean of the faculty of applied mathematics and computer science.
orlovich@bsu.by
Mikita A. Shutro, student at the faculty of applied mathematics and computer science.
fpm.shutro@bsu.by

Introduction

In this paper, we use graph-theoretic terminology of J. Bondy and U. Murty [1] (unless noted otherwise), and computational complexity terminology of M. Garey and D. Johnson [2]. For concepts related to domination in graphs, we follow T. Haynes et al. [3].

Throughout this paper G is a simple (i. e. a finite, undirected, loopless and without multiple edges) graph with the vertex set $V(G)$ and the edge set $E(G)$. A class of graphs is a set of graphs closed under isomorphism. A graph is trivial if it has only one vertex. A graph with no edges is called empty. An edge of G is two-element subset of $V(G)$. Often in the paper, an edge $\{u, v\}$ will be written uv or vu , dropping the curly brackets. The vertices u and v are called end-vertices of the edge uv .

Let G be a graph. For any vertex $v \in V(G)$, the open neighbourhood of v , denoted by $N(v)$, is defined by $\{u \in V(G) \mid uv \in E(G)\}$. The closed neighbourhood of v , denoted by $N[v]$, is the set $N(v) \cup \{v\}$. The neighbourhood of an edge $e = uv \in E(G)$ is defined as $N(e) = N(u) \cup N(v)$. Clearly, $\{u, v\} \subseteq N(e)$. The degree of v is defined as $|N(v)|$. A vertex of degree 0 is called isolated. We say that a graph is isolate-free if it has no isolated vertices.

As usual, by P_n we denote the chordless path on n vertices, by C_n the chordless cycle on n vertices, and by $K_{1,n}$ the star on $n+1$ vertices, i. e. the complete bipartite graph with partition classes of cardinalities 1 and n . Given two graphs – G_1 and G_2 , we denote by $G_1 \cup G_2$ the disjoint union of G_1 and G_2 . In particular, for a positive integer n , the disjoint union of n copies of a graph G is denoted by nG . The complement of a graph G is denoted by $\bar{G}$.

For a graph G and a non-empty subset S of $V(G)$, the subgraph $G(S)$ of G induced by S has S as its vertex set and two vertices – u and v – are adjacent in $G(S)$ if and only if u and v are adjacent in G . A subgraph H of a graph G is called an induced subgraph if there is a non-empty subset S of $V(G)$ such that $H = G(S)$. For a (finite or infinite) set M of graphs, a graph G is called M -free if no induced subgraph of G is isomorphic to a graph in M .

As usual, the term «minimal» («maximal», respectively) with respect to the sets having some property means an inclusion-wise minimal (an inclusion-wise maximal, respectively) set with this property, while the terms «minimum» and «maximum» refer to the cardinality of a set with a given property.

A subset $S \subseteq V(G)$ is an independent set of G if no two vertices in S are adjacent. A subset $D \subseteq V(G)$ is a dominating set of G if each vertex of $V(G) \setminus D$ is adjacent to a vertex of D . If, in addition, D is an independent set, then D is an independent dominating set. The domination number $\gamma(G)$ of G is the minimum cardinality of a dominating set in G . The independent domination number $i(G)$ of G is the minimum cardinality of an independent dominating set in G , or equivalently, is the minimum cardinality of a maximal independent set in G . The independence number $\alpha(G)$ of G is the maximum cardinality of an independent set in G .

It is well known that for any graph G , $\gamma(G) \leq i(G) \leq \alpha(G)$. A graph is well-covered if all its maximal independent sets have the same cardinality, equivalently if every maximal independent set is maximum. In other words, a graph G is well-covered, if and only if $i(G) = \alpha(G)$ is satisfied. The concept of a well-covered graph was introduced by M. Plummer [4] in 1970. These graphs are of interest because, whereas the independent set problem is NP-complete for general graphs, the maximum independent set can be found efficiently in a well-covered graph by using a simple greedy algorithm.

The problem of recognising well-covered graphs is known to be co-NP-complete [5; 6]. It remains co-NP-complete even for $K_{1,4}$ -free graphs [7], for circulant graphs [8], and for $\{2P_5, K_{1,10}\}$ -free squares of line graphs [9]. However, the problem is solvable in polynomial time for $K_{1,3}$ -free graphs [10; 11], for graphs with girth at least 5 [12], for graphs that contain neither C_4 nor C_5 as a subgraph [13], for graphs with a bounded maximum degree [14], or for chordal graphs [15].

The following two variants of domination in graphs, which are called vertex-edge domination and edge-vertex domination, were introduced by K. Peters in work [16] and further studied in papers [17–19]. A vertex $u \in V(G)$ is said to ve -dominate an edge $xy \in E(G)$ if either (1) $u = x$ or $u = y$, that is, u is incident to xy or (2) ux or uy is an edge in G , that is, u is incident to an edge that is adjacent to xy . In other words, a vertex u ve -dominates the edges incident to vertices in $N[u]$. An edge $uw \in E(G)$ is said to ev -dominate a vertex $x \in V(G)$ if either (1) $u = x$ or $w = x$, that is, uw is incident to x or (2) ux or wx is an edge in G , that is, x is adjacent to u or w . In other words, an edge uw ev -dominates the vertices in $N(uw)$.

We need some additional concepts about private edges and private vertices (see [20–22]). Let G be a graph and let S be a subset of the vertex set $V(G)$ of G . A vertex $u \in S$ has a private edge $xy \in E(G)$ (with respect to S), if (1) u ve -dominates the edge xy (i. e. u is incident to xy or u is adjacent to either x or y), and (2) no other vertex in S ve -dominates xy (i. e. for all vertices $w \in S \setminus \{u\}$, w is not incident to xy and w is not adjacent to either x or y). It was shown in papers [20; 21] that a ve -dominating set S of a non-trivial connected graph G is minimal if and only if each vertex $u \in S$ has at least one private edge (with respect to S).

Now let F be a subset of the edge set $E(G)$ of G . An edge $uw \in F$ has a private vertex $x \in V(G)$ (with respect to F), if (1) uw ev -dominates the vertex x (i. e. uw is incident to x or x is adjacent to either u or w), and (2) no other edge in F ev -dominates w (i. e. for all edges $ab \in F \setminus \{uw\}$, ab is neither incident to x and neither a nor b is adjacent to x). It was shown in paper [22] that an ev -dominating set F of a non-trivial connected graph G is minimal if and only if each edge $uw \in F$ has at least one private vertex (with respect to F).

The vertex-edge domination number (or simply the ve -domination number) $\gamma_{ve}(G)$ is the minimum cardinality of a ve -dominating set of G . An independent vertex-edge dominating set (or simply independent ve -dominating set) is a vertex subset that is both independent and ve -dominating. The independent vertex-edge domination number (or simple independent ve -domination number) $i_{ve}(G)$ is the minimum cardinality of an independent ve -dominating set, and the upper independent vertex-edge domination number (or simply upper independent ve -domination number) $\alpha_{ve}(G)$ is the maximum cardinality taken over all minimal independent ve -dominating sets of G .

The edge-vertex domination number (or simply the ev -domination number) $\gamma_{ev}(G)$ is the minimum cardinality of an ev -dominating set of G . An independent edge-vertex dominating set (or simply independent ev -dominating set) is an edge subset that is both independent (i. e. a matching) and ev -dominating. The independent edge-vertex domination number (or simple independent ev -domination number) $i_{ev}(G)$ is the minimum cardinality of an independent ev -dominating set, and the upper independent edge-vertex domination number (or simply upper independent ev -domination number) $\alpha_{ev}(G)$ is the maximum cardinality taken over all minimal independent ev -dominating sets of G .

For any graph G , the above vertex-edge domination parameters are related by the following chain of inequalities established in paper [17]: $\gamma_{ve}(G) \leq i_{ve}(G) \leq \alpha_{ve}(G)$. Let us remark that the empty graph O_n on $n \geq 1$ vertices has a unique minimal ve -dominating set, namely, the empty set. So, $\gamma_{ve}(O_n) = i_{ve}(O_n) = \alpha_{ve}(O_n) = 0$.

Note that only graphs without isolated vertices have ev -dominating sets. It is well-known (see [17]) that if G is an isolate-free graph, then $\gamma_{ev}(G) \leq i_{ev}(G) \leq \alpha_{ev}(G)$. Moreover, as shown in paper [17], the equality $\gamma_{ev}(G) = i_{ev}(G)$ between edge-vertex domination parameters $\gamma_{ev}(G)$ and $i_{ev}(G)$ holds for all isolate-free graphs G .

A graph G is well ve -covered if $i_{ve}(G) = \alpha_{ve}(G)$, i. e. all its minimal independent ve -dominating sets have the same cardinality $\alpha_{ve}(G)$. This concept was introduced by R. Boutrig et al. [21]. They proved that the problem of recognising well ve -covered graphs is co-NP-complete and presented a constructive characterisation of well ve -covered trees.

In order to characterise the trees T with $i_{ev}(T) = \alpha_{ev}(T)$, R. Boutrig and M. Chellali [22] introduced the concept of the well ev -covered graph, which is similar to the concept of well ve -covered graph but in terms of the independent ev -domination. A graph G is well ev -covered if $i_{ev}(G) = \alpha_{ev}(G)$, i. e. all its minimal independent ev -dominating sets have the same cardinality $\alpha_{ev}(G)$. In paper [22], R. Boutrig and M. Chellali proposed some open problems concerning well ev -covered graphs, among them:

1. Can you recognise a well ev -covered graph in polynomial time?
2. Characterise the well ev -covered graphs.

It has been shown by M. Chellali, in his recent paper [23], that the problem of recognising well ev -covered graphs is co-NP-complete. This result shows that the property of being not well ev -covered is NP-complete. Thus, returning to the open problems formulated in article [22], one can note that the existence of a good characterisation of the well ev -covered graphs is very unlikely.

The remaining part of the present paper is organised as follows.

In the section «Well ve -covered graphs», we prove that the problem of recognising well ve -covered graphs is co-NP-complete even when restricted to the class of $K_{1,4}$ -free graphs. This result strengthens the result given in paper [21], where it is shown that the problem is co-NP-complete for general graphs. We also prove that, under the same restriction, the problem of recognising a graph that has minimal independent ve -dominating sets of at most t distinct cardinalities is co-NP-complete for any given integer $t \geq 2$.

The section «Well ev -covered graphs» contains the proof that the problem of recognising well ev -covered graphs is still co-NP-complete even for some graph classes defined by certain forbidden induced subgraphs. Note that our polynomial transformation in the proof of this result is simpler than that in paper [23] and allow us to reduce the problem from general graphs to connected graphs in a proper subclass of $\{C_4, \overline{K_1 \cup C_4}, K_{1,5}, 2P_4\}$ -free graphs.

In the section «Hereditary properties», we show that the classes of well ve -covered graphs and well ev -covered graphs are not hereditary, and therefore, they have no characterisations in terms of forbidden induced subgraphs. As an alternative, we provide finite characterisations for the maximal hereditary subclasses of these classes in terms of forbidden induced subgraphs.

In the section «Related results», we apply the previous results to obtain new cases where the decision problems associated with the computation of $i_{ve}(G)$, $i_{ev}(G)$ and $\gamma_{ev}(G)$ are NP-complete. Also in this section we introduce a concept of a well paired-dominated graph and study the computational complexity of recognising such graphs.

The section «Conclusions» provides a summary of our results and presents some open problems.

In what follows all decision problems are stated in the standard instance – question format [2].

Well ve -covered graphs

In this section we shall consider the following decision problem.

Non-Well ve -Covered Graph problem

Instance: a graph G .

Question: is it true that the graph G is not well ve -covered? In other words, are there two minimal independent ve -dominating sets of distinct cardinalities in G ?

We first prove that the Non-Well ve -Covered Graph problem is an NP-complete problem and thereby establish that recognising the class of well ve -covered graphs is a co-NP-complete problem. Then we extend this result by showing that recognising a graph that has minimal independent ve -dominating sets of at most t distinct cardinalities is a co-NP-complete problem for any $t \geq 2$. In our proof, we use a transformation from the following well-known NP-complete decision problem (see [2]).

Three-Dimensional Matching (3-DM) problem

Instance: pairwise disjoint sets X, Y, Z of some finite cardinality k , and a collection M of three-element sets, where each member of M includes exactly one element from each of X, Y and Z .

Question: is there a set of pairwise disjoint members of M , whose union is $X \cup Y \cup Z$?

We say that the instance of 3-DM problem is solvable if there exists a set of pairwise disjoint members of M , whose union is $X \cup Y \cup Z$.

Theorem 1. *The Non-Well ve -Covered Graph problem is NP-complete for $K_{1,4}$ -free graphs.*

Proof. It can be checked in polynomial time whether a ve -dominating set is minimal, since the property of being a ve -dominating set is superhereditary (for more details see [3]). Thus, it is obvious that the Non-Well ve -Covered Graph problem is a member of NP since we can guess subsets $D_1, D_2 \subseteq V(G)$ and, in polynomial time, verify that D_1 and D_2 are minimal independent ve -dominating sets of distinct cardinalities.

To establish NP-completeness we start by describing how to construct an instance of the Non-Well ve -Covered Graph problem from one for the 3-DM problem. Let $I = (X, Y, Z, M)$ be an instance of the 3-DM problem, where $X = \{x_1, x_2, \dots, x_k\}$ and $Y = \{y_1, y_2, \dots, y_k\}$. Let $\{a, b_i, c_i \mid i=1, 2, \dots, k\}$ be a set of $2k+1$ new elements, each of which does not belong to the union $X \cup Y \cup Z$. We define $M_x = \{\{x_i, a\} \mid i=1, 2, \dots, k\}$ and $M_y = \{\{y_i, b_i\}, \{y_i, c_i\} \mid i=1, 2, \dots, k\}$. Let $G(I)$ be the intersection graph of the subset system $S = M \cup M_x \cup M_y \cup \{\{a\}\}$, i. e. the vertex set of $G(I)$ is S , and two vertices are adjacent if and only if the corresponding subsets have a non-empty intersection. It is easy to see that the construction of the graph $G(I)$ can be accomplished in polynomial time.

An example of graph $G(I)$ constructed for $k=3$ and

$$M = \{\{x_1, y_2, z_1\}, \{x_3, y_2, z_3\}, \{x_2, y_1, z_1\}, \{x_1, y_2, z_3\}, \{x_3, y_1, z_2\}, \{x_2, y_3, z_1\}\}$$

is given in fig. 1, where (p, q, r) stands for the vertex $\{x_p, y_q, z_r\}$.

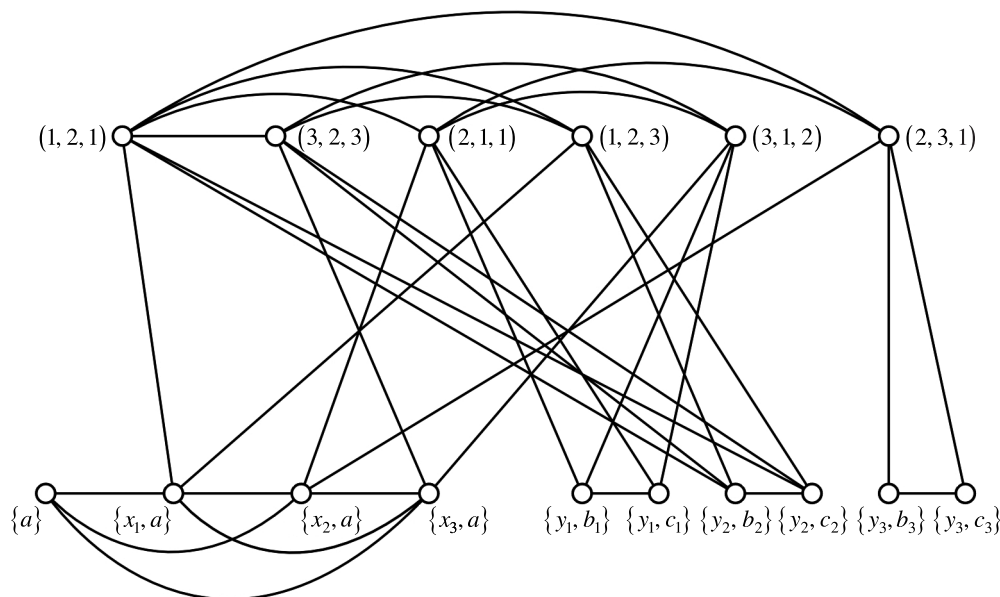


Fig. 1. An example of the graph $G(I)$

Note that the graph $G(I)$ is connected and has always a minimal independent ve -dominating set of cardinality $k + 1$, for example $\{\{a\}\} \cup \{\{y_i, b_i\} \mid i = 1, 2, \dots, k\}$ is such a set.

To prove the theorem, it is sufficient to show that the following three claims are valid.

Claim 1. The graph $G(I)$ is $K_{1,4}$ -free.

Proof. Indeed, since $|T| \leq 3$ for each $T \in S$, any four sets, each intersecting with T , cannot be pairwise disjoint. This, in turn, means that the graph $G(I)$ cannot contain a star $K_{1,4}$ centred at vertex T as an induced subgraph. The claim is proved.

Claim 2. For any minimal independent ve -dominating set D of $G(I)$, the inequalities $k \leq |D| \leq k + 1$ hold.

Proof. Let D be an arbitrary minimal independent ve -dominating set of $G(I)$. Since any $k + 2$ members of the subset system S obviously contain an intersecting pair, we have that any independent (in particular, minimal) ve -dominating set of $G(I)$ contains at most $k + 1$ vertices, and hence $|D| \leq k + 1$.

For $i = 1, 2, \dots, k$, define

$$S_i = \{\{y_i, b_i\}, \{y_i, c_i\}, \{x_p, y_i, z_r\} \mid \{x_p, y_i, z_r\} \in M, 1 \leq p, r \leq k\}.$$

Let e_i be the edge of the graph $G(I)$ connecting vertices $\{y_i, b_i\}$ and $\{y_i, c_i\}$, for $1 \leq i \leq k$. Clearly, the edge e_i can only be ve -dominated by vertices from S_i . Since the set S_i is a clique of the graph $G(I)$, we have that $|D \cap S_i| = 1$, for each $i = 1, 2, \dots, k$. Thus, $|D| \geq k$, and the claim is proved.

Claim 3. The instance $I = (X, Y, Z, M)$ is solvable for the 3-DM problem if and only if the graph $G(I)$ has a minimal independent ve -dominating set of cardinality k .

Proof. Let the instance $I = (X, Y, Z, M)$ is solvable for the 3-DM problem, i. e. M contains a subset M' of k pairwise disjoint members, whose union is equal to $X \cup Y \cup Z$. Construct a set $D \subseteq V(G(I))$ of cardinality k by including in it all the vertices of the graph $G(I)$ that correspond to elements of M' . It is easy to see that D is independent and ve -dominating set of $G(I)$. The minimality of the set D follows from the fact that every vertex $\{x_p, y_q, z_r\} \in D$ has the edge $\{\{y_q, b_q\}, \{y_q, c_q\}\} \in E(G(I))$ as a private edge (with respect to D).

Assume now that the graph $G(I)$ has a minimal independent ve -dominating set D of cardinality k . For $1 \leq i \leq k$, let e_i be the edge of $G(I)$ between the vertices $\{y_i, b_i\}$ and $\{y_i, c_i\}$, and let S_i be the set defined as in the proof of claim 2. Then as in the proof of claim 2, we have for each $i = 1, 2, \dots, k$ that $|D \cap S_i| = 1$, and therefore, the intersection of D with the set $M_x \cup \{\{a\}\}$ is empty. Suppose that $\{y_i, b_i\} \in D$ for some $i \in \{1, 2, \dots, k\}$. Then $|D \cap M| < k$, and hence there exists an element $x_j \in X$, which does not belong to any member of $D \cap M$. But then no vertex in D , which ve -dominates the edge $\{\{x_j, a\}, \{a\}\} \in E(G(I))$, a contradiction since D is a ve -dominating set

of $G(I)$. Thus, the intersection of D with the set M_y is also empty, and consequently $D \subseteq M$. Since D is an independent ve -dominating set its vertices correspond to exactly k pairwise disjoint members of M , whose union is $X \cup Y \cup Z$. Therefore, the instance I is solvable. This completes the proof of the claim.

So, we have obtained a polynomial transformation from the 3-DM problem to the Non-Well ve -Covered Graph problem within $K_{1,4}$ -free graphs, and the theorem is proved.

Let $W_{ve}(t)$ be the class of graphs, each of which has minimal independent ve -dominating sets of at most t cardinalities. Note that $W_{ve}(1)$ is the class of well ve -covered graphs and $W_{ve}(t) \subset W_{ve}(t+1)$ holds for $t \geq 1$. Consider the following decision problem.

Non- t -Well ve -Covered Graph problem

Instance: a graph G and an integer number $t \geq 2$.

Question: is it true that $G \notin W_{ve}(t)$? In other words, are there $t+1$ minimal independent ve -dominating sets of distinct cardinalities in G ?

Based on theorem 1, the following result is obtained.

Theorem 2. For any $t \geq 2$, the Non- t -Well ve -Covered Graph problem is NP-complete even for $K_{1,4}$ -free graphs.

Proof. It is not hard to see that the Non- t -Well ve -Covered Graph problem belongs to NP. To prove the NP-completeness, we will use a polynomial transformation from the Non-Well ve -Covered Graph problem, which is NP-complete by theorem 1. Let G be a non-trivial connected $K_{1,4}$ -free graph, which constitutes an instance for the Non-Well ve -Covered Graph problem. Consider the graph H , which is the disjoint union of t components $G_1, G_2, \dots, G_t$, each of which is isomorphic to the graph G . It is easy to see that the graph H is $K_{1,4}$ -free, and that H can be constructed in polynomial time in the number of vertices of G . To complete the proof of the theorem, it suffices to show the following.

Claim 4. $G \in W_{ve}(1)$ if and only if $H \in W_{ve}(t)$.

Proof. Assume that G is a well ve -covered graph, i. e. $G \in W_{ve}(1)$. Let D_1 and D_2 be any two minimal independent ve -dominating sets of the graph H . The intersection of $D_i, i \in \{1, 2\}$, with the vertex set $V(G_j)$ of the j^{th} component G_j in H is a minimal independent ve -dominating set of the graph G_j . Since G_j is isomorphic to the graph G and all minimal independent ve -dominating sets in G have the same cardinality, we conclude that $|D_1| = |D_2|$. Hence, $H \in W_{ve}(1)$, and since $W_{ve}(1) \subset W_{ve}(t)$ for $t \geq 2$, we obtain $H \in W_{ve}(t)$.

For the converse implication assume that $H \in W_{ve}(t)$ for some $t \geq 2$. Suppose on the contrary that the graph G is not well ve -covered. Then there are minimal independent ve -dominating sets D_1 and D_2 in G with $|D_1| \neq |D_2|$. Without loss of generality let $|D_1| < |D_2|$. For $i = 0, 1, \dots, t$, define the set S_i by

$$S_i = (D_1^1 \cup D_1^2 \cup \dots \cup D_1^i) \cup (D_2^{i+1} \cup D_2^{i+2} \cup \dots \cup D_2^t),$$

where D_1^j and D_2^j are copies of the sets D_1 and D_2 in the j^{th} component G_j of the graph H , respectively. Clearly, for each $i \in \{0, 1, \dots, t\}$, the set S_i is a minimal independent ve -dominating set of H and $|S_i| = i|D_1| + (t-i)|D_2|$. Therefore, for $0 \leq k < l \leq t$, we obtain $|S_k| - |S_l| = (|D_2| - |D_1|)(l-k)$ and, consequently, $|S_0| > |S_1| > \dots > |S_t|$. But this contradicts the fact that $H \in W_{ve}(t)$, and the claim is proved.

So, we have obtained a polynomial transformation from the Non-Well ve -Covered Graph problem restricted to $K_{1,4}$ -free graphs to the Non- t -Well ve -Covered Graph problem restricted to $K_{1,4}$ -free graphs, and the theorem is proved.

Well ev -covered graphs

In this section we shall consider the following decision problem.

Non-Well ev -Covered Graph problem

Instance: a graph G .

Question: is it true that the graph G is not well ev -covered? In other words, are there two minimal independent ev -dominating sets of distinct cardinalities in G ?

We prove that the Non-Well ev -Covered Graph problem is an NP-complete problem and thereby establish that recognising the class of well ev -covered graphs is a co-NP-complete problem. In our proof, we use a polynomial transformation from the following well-known NP-complete decision problem (see [2]).

Three-Satisfiability (3Sat) problem

Instance: a collection $C = \{c_1, c_2, \dots, c_m\}$ of m clauses over a set $X = \{x_1, x_2, \dots, x_n\}$ of n Boolean variables such that $|c_j| = 3$ for $j = 1, 2, \dots, m$.

Question: is there a truth assignment for X that simultaneously satisfies all the clauses in C ?

Recall that a truth assignment for X is a mapping $\varphi : X \rightarrow \{0, 1\}$ that assigns a value $\varphi(x_i) \in \{0, 1\}$ to each variable $x_i \in X$. We extend φ to the set $X \cup \{\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n\}$ of the literals over X by setting $\varphi(\bar{x}_i) = \overline{\varphi(x_i)}$, $i = 1, 2, \dots, n$. A literal $\ell \in X \cup \{\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n\}$ is true under φ if $\varphi(\ell) = 1$. A clause over X is a disjunction of some literals. A truth assignment φ for X satisfies a clause $c_j \in C$ if c_j involves at least one true literal under φ . A collection C of clauses over X is satisfiable if and only if there exists some truth assignment for X that simultaneously satisfies all the clauses in C .

In order to formulate the following result, we need to introduce special notations for some particular classes of graphs: A is the class of $\{K_1 \cup C_4, K_{1,5}, 2P_4\}$ -free graphs, and B is the class of $\{C_4, C_6, C_7, C_8, C_9\}$ -free graphs. In fact, it is not hard to see that any graph in $A \cap B$ is also C_p -free and $\bar{C}_q$ -free for each $p \geq 10$ and $q \geq 6$, as well as P_s -free and $\bar{P}_t$ -free for each $s \geq 9$ and $t \geq 5$.

Theorem 3. *The Non-Well ev -Covered Graph problem is NP-complete in the class of graphs $A \cap B$.*

Proof. It can be checked in polynomial time whether an ev -dominating set is minimal, since the property of being an ev -dominating set is superhereditary (for more details see [3]). Thus, the Non-Well ev -Covered Graph problem is in NP since we can guess subsets $E_1, E_2 \subseteq E(G)$ and, in polynomial time, verify that E_1 and E_2 are minimal independent ev -dominating sets of distinct cardinalities. To show that it is NP-complete, we construct a polynomial transformation from the 3Sat problem. Without loss of generality, we may assume that each variable appears (as x or $\bar{x}$) in at least one clause and no clause in C contains a variable x_i and its negation $\bar{x}_i$ because such a clause is satisfied by any truth assignment and can be eliminated. We also assume that all clauses are distinct and $m = |C| \geq 2$. The proof uses ideas from paper [9].

Given an instance $J = (C, X)$ of the 3Sat problem, construct the following graph $G(J)$. Some vertices of this graph are associated with the parameters of the 3Sat problem and use the same notation, and some vertices are auxiliary. Specifically, the vertex set of the graph $G(J)$ is the union $C' \cup X'$, where $C' = \{c_j \mid j = 1, 2, \dots, m\}$ and $X' = \{x_i, y_i, z_i, \bar{x}_i \mid i = 1, 2, \dots, n\}$. Edges of $G(J)$ are defined as follows.

1. The set C' induces a clique.
2. The set X' induces n vertex-disjoint stars $K_{1,3}^i$, $i = 1, 2, \dots, n$, each one with four vertices x_i, y_i, z_i and $\bar{x}_i$ and the edges $x_i y_i, y_i z_i$ and $y_i \bar{x}_i$.
3. For each clause $c_j = (\ell_j^1 \vee \ell_j^2 \vee \ell_j^3)$, $j = 1, 2, \dots, m$, introduce the three edges $c_j \ell_j^1, c_j \ell_j^2$ and $c_j \ell_j^3$ between C' and X' in $G(J)$.

The graph $G(J)$ can obviously be constructed in time polynomial in m and n . Figure 2 shows an example of graph $G(J)$ when $X = \{x_1, x_2, x_3, x_4\}$ and $C = \{c_1, c_2, c_3, c_4, c_5\}$, where $c_1 = (x_1 \vee x_2 \vee x_3)$, $c_2 = (\bar{x}_1 \vee x_2 \vee x_3)$, $c_3 = (x_1 \vee x_3 \vee \bar{x}_4)$, $c_4 = (\bar{x}_2 \vee \bar{x}_3 \vee x_4)$, $c_5 = (\bar{x}_1 \vee \bar{x}_3 \vee x_4)$.

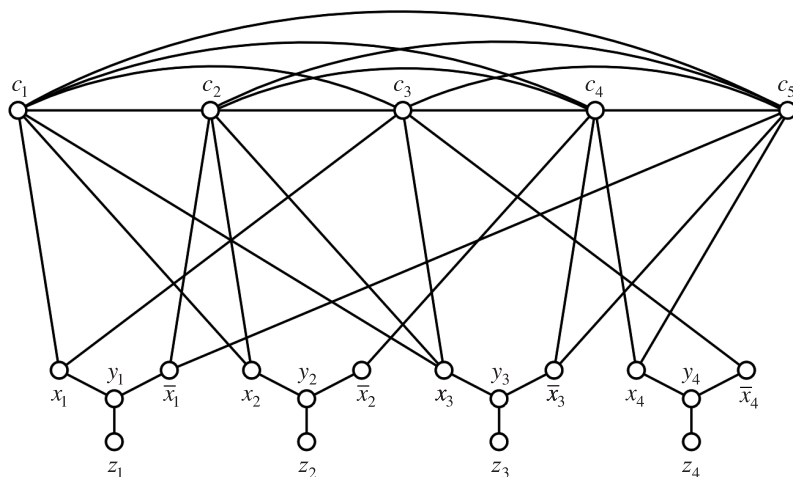


Fig. 2. An example of the graph $G(J)$

Note that the graph $G(J)$ is connected and has always a minimal independent ev -dominating set of cardinality $n + 1$, for example $\{c_1c_2\} \cup \{y_iz_i \mid i = 1, 2, \dots, n\}$ is such a set. Recall that $m \geq 2$.

To prove the theorem, it is sufficient to show that the following three claims are valid.

Claim 5. The graph $G(J)$ belongs to the class $A \cap B$.

Proof. It is not difficult to show that the graph $G(J)$ does not contain any graph in $\{\overline{K_1 \cup C_4}, K_{1,5}, 2P_4\} \cup \{C_4, C_6, C_7, C_8, C_9\}$ as an induced subgraph. Hence, $G(J)$ is in the class $A \cap B$, as required. The claim is proved.

Let us denote $X_i = \{x_iz_i, y_iz_i, y_i\bar{x}_i\}$, $i = 1, 2, \dots, n$.

Claim 6. For any minimal independent ev -dominating set E' in $G(J)$, the following statements hold:

- (i) $|E' \cap X_i| = 1$ for $i = 1, 2, \dots, n$;
- (ii) $|E' \cap EN(C')| \leq 1$, where $EN(C')$ is the set of all edges in $G(J)$ that have at least one end-vertex in C' ;
- (iii) $|E'| \in \{n, n + 1\}$.

Proof. (i) If $E' \cap X_i = \emptyset$ for some $i \in \{1, 2, \dots, n\}$, then the vertex z_i cannot be ev -dominated. Therefore, $|E' \cap X_i| \geq 1$ for each $i = 1, 2, \dots, n$. On the other hand, since E' is an independent set of edges and, for each $i = 1, 2, \dots, n$, the set X_i forms the star $K_{1,3}$ in $G(J)$, we have $|E' \cap X_i| \leq 1$, $i = 1, 2, \dots, n$. Thus, $|E' \cap X_i| = 1$ for $i = 1, 2, \dots, n$.

(ii) Suppose on the contrary that the intersection $E' \cap EN(C')$ contains two edges – e_1 and e_2 . Let u be a private vertex (with respect to E') of the edge e_1 . Suppose that $u \in X'$, say $u \in V(K_{1,3}^i)$ for some $i \in \{1, 2, \dots, n\}$. Then by the statement (i) there exists the edge $e' \in E' \cap X_i$ such that e' is distinct from e_1 , and e' ev -dominates u , which is impossible. Hence, $u \in C'$. Finally, since C' is a clique in $G(J)$ and $e_2 \in E' \cap EN(C')$, we conclude that e_2 ev -dominates u , again a contradiction. This contradiction shows that $|E' \cap EN(C')| \leq 1$, as required.

(iii) By the statement (i), the set E' contains exactly n edges of $G(X')$. Thus, $|E'| \geq n$. The fact that $|E'| \leq n + 1$ follows from the statement (ii).

The claim is proved.

Claim 7. C is satisfiable if and only if the graph $G(J)$ has a minimal independent ev -dominating set of cardinality n .

Proof. First, suppose that there exists a truth assignment ϕ satisfying C . We construct a set $E' \subset \bigcup_{i=1}^n X_i$ by choosing edges that correspond to the true literals. Namely, for each $i = 1, 2, \dots, n$, do the following. If $\phi(x_i) = 1$, then the edge x_iz_i is included in E' . On the other hand, if $\phi(x_i) = 0$, then the edge $y_i\bar{x}_i$ is included in E' . Thus, E' is of cardinality n .

We shall prove that E' is a minimal independent ev -dominating set in $G(J)$. Indeed, since $X_i \cap X_j = \emptyset$ (for $1 \leq i \neq j \leq n$), we have that E' is a matching, i. e. an independent set of edges. Since ϕ satisfies C , each vertex of C' is adjacent to an end-vertex of an edge in E' , and hence is ev -dominated by such an edge. Since for each $i = 1, 2, \dots, n$ the set X_i form a star $K_{1,3}$ in $G(J)$, we have that each vertex of X' is either incident to an edge in E' or is adjacent to an end-vertex of an edge in E' , and hence is ev -dominated by such an edge. Therefore, E' is an ev -dominating set in $G(J)$. Finally, the minimality of the set E' follows directly from the fact that for each $i = 1, 2, \dots, n$ the vertex z_i is a private vertex of the unique edge in E' that belongs to the set X_i .

Conversely, assume that E' is a minimal independent ev -dominating set of cardinality n in the graph $G(J)$. By claim 6 (i), we have that $|E' \cap X_i| = 1$ for $i = 1, 2, \dots, n$. Therefore, $E' \subset E(G(X'))$ and E' covers at most one of the vertices $x_i, \bar{x}_i$ for each $i = 1, 2, \dots, n$. Since E' ev -dominates every vertex in C' , we can define a partial truth assignment ϕ' satisfying C by deciding a literal $\ell_i \in \{x_i, \bar{x}_i\}$ to be true if the corresponding vertex of C' is ev -dominated by the edge $\ell_i y_i \in E' \cap X_i$. It remains to extend, if necessary, ϕ' to a full assignment. This completes the proof of the claim.

So, we have obtained a polynomial transformation from the 3Sat problem to the Non-Well ev -Covered Graph problem restricted to the class of graphs $A \cap B$, and the theorem is proved.

Let $W_{ev}(t)$ be the class of graphs, each of which has minimal independent ev -dominating sets of at most t cardinalities. Note that $W_{ev}(1)$ is the class of well ev -covered graphs and $W_{ev}(t) \subset W_{ev}(t + 1)$ holds for $t \geq 1$. Consider the following decision problem.

Non- t -Well ev -Covered Graph problem

Instance: a graph G and an integer number $t \geq 2$.

Question: is it true that $G \notin W_{ev}(t)$? In other words, are there $t + 1$ minimal independent ev -dominating sets of distinct cardinalities in G ?

The following result can be obtained by applying theorem 3 and the same approach that we have used to prove theorem 2.

Theorem 4. For any $t \geq 2$, the Non- t -Well ev -Covered Graph problem is NP-complete even in the class of graphs $A \cap B$.

Hereditary properties

Let us remind basic notions related to hereditary properties of graphs referring for more details to the monograph by S. Kitaev and V. Lozin (see [24, chap. 2]).

A class of graphs is called hereditary if every induced subgraph of a graph in this class also belongs to the class. The family of hereditary classes includes many important classes of graphs, for instance, bipartite graphs, chordal graphs, planar graphs, line graphs, graphs of vertex degree at most k (for a fixed k), etc.

It is well known that a class X of graphs is hereditary if and only if X is M -free (i. e. every $G \in X$ is M -free) for some set M . In this case graphs in M are called forbidden induced subgraphs for the class X . Generally, for a hereditary class X the set M of forbidden induced subgraphs is not unique. For instance, M can be defined as the set of all graphs that are not in X . However, to describe X it is sufficient to find only so-called minimal forbidden induced subgraphs. A forbidden induced subgraph H for hereditary class X is minimal if H is not in X but every proper induced subgraph of H is in X . It is known that for every hereditary class the set of minimal forbidden induced subgraphs exists and is unique.

As noted in paper [25], every (hereditary or not) class X of graphs can be «approximated» by two hereditary classes as follows: by $\lfloor X \rfloor$ we denote the maximal hereditary subclass contained in X and by $\lceil X \rceil$ the minimal hereditary class containing X . It is known that $\lfloor X \rfloor$ and $\lceil X \rceil$ are uniquely defined and that $\lfloor X \rfloor \subseteq X \subseteq \lceil X \rceil$ with equalities holding if and only if X is a hereditary class.

In order to prove that the classes $W_{ve}(1)$ and $W_{ev}(1)$ of well ve -covered graphs and well ev -covered graphs are not hereditary classes, let us prove the following result.

Theorem 5. *For any graph H , there exists a graph G that contains H as an induced subgraph and G is both a well ve -covered graph and a well ev -covered graph.*

Proof. Given a graph H with the vertex set $V(H) = \{u_1, u_2, \dots, u_n\}$, $n \geq 1$, construct a new graph G as follows. For each vertex u_i , $i = 1, 2, \dots, n$, add two new vertices – a_i and b_i , and join u_i to a_i and a_i to b_i , respectively. More precisely,

$$V(G) = V(H) \cup \{a_i, b_i \mid i = 1, 2, \dots, n\} \text{ and } E(G) = E(H) \cup \{u_i a_i, a_i b_i \mid i = 1, 2, \dots, n\}.$$

Clearly, by the construction, H is an induced subgraph of G .

Let D be any minimal independent ve -dominating set of G , and let $U = D \cap V(H)$. Then for each vertex $u_i \in V(H) \setminus U$, the set D contains exactly one of the two vertices – a_i or b_i , since otherwise the edge $a_i b_i$ cannot be ve -dominated. Therefore, $|D| = |U| + |V(H) \setminus U| = n$, and hence G is well ve -covered.

Now let F be any minimal independent ev -dominating set of G . It is easy to see that for each $i = 1, 2, \dots, n$, the set F contains exactly one of the two edges – $u_i a_i$ or $a_i b_i$, since otherwise the vertex b_i cannot be ev -dominated. By minimality of F , we have $F \cap E(H) = \emptyset$. Therefore, $|F| = n$, and hence G is well ev -covered.

The theorem is proved.

By theorem 5 we conclude that both hereditary classes – $\lfloor W_{ve}(1) \rfloor$ and $\lceil W_{ev}(1) \rceil$ – coincide with the class of all graphs.

Now we consider a hereditary subclass of the class $W_{ve}(1)$. A graph G is called perfectly well ve -covered if every induced subgraph of G is well ve -covered. We characterise perfectly well ve -covered graphs in terms of minimal forbidden induced subgraphs.

In the sequel of this section, the notation $y \sim z$ ($y \not\sim z$, respectively) means that the vertices y and z are adjacent (non-adjacent, respectively). For disjoint sets of vertices Y and Z , the notation $Y \sim Z$ ($Y \not\sim Z$, respectively) means that every vertex of Y is adjacent (non-adjacent, respectively) to every vertex of Z . In the case when $Y = \{y\}$, we also write $y \sim Z$ and $y \not\sim Z$ instead of $\{y\} \sim Z$ and $\{y\} \not\sim Z$, respectively. Also, for the sake of simplicity, a subgraph of G induced by a vertex set $\{u_1, u_2, \dots, u_k\} \subseteq V(G)$ will be denoted by $G(u_1, u_2, \dots, u_k)$ instead of $G(\{u_1, u_2, \dots, u_k\})$.

Theorem 6. *For a graph G , the following statements are equivalent:*

- (i) G is a perfectly well ve -covered graph;
- (ii) G is a $\{G_1, G_2\}$ -free graph, where the graphs G_1 and G_2 are shown in fig. 3.

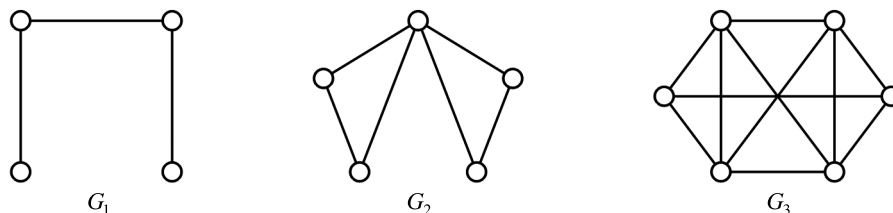


Fig. 3. The graphs G_1 , G_2 and G_3

Proof. We prove first the implication (i) $\Rightarrow$ (ii). It is easy to check that $i_{ve}(H)=1$ and $\alpha_{ve}(H)=2$ for $H \in \{G_1, G_2\}$. Therefore, both graphs – G_1 and G_2 – are not perfectly well ve -covered graphs, and they are minimal. It follows that none of them can be an induced subgraph of a perfectly well ve -covered graph.

Now we prove the converse implication (ii) $\Rightarrow$ (i). Suppose that G is a $\{G_1, G_2\}$ -free graph, but G is not a perfectly well ve -covered graph. We may assume that G is minimal, i. e. each proper induced subgraph of G is perfectly well ve -covered. By definition, there are two minimal independent ve -dominating sets – D_1 and D_2 – in G with $|D_1| \neq |D_2|$. Without loss of generality, we may assume that $|D_1| > |D_2|$. Since D_1, D_2 are minimal ve -dominating sets and $|D_1 \setminus D_2| > |D_2 \setminus D_1|$, there are three distinct vertices – $u_1, u_2 \in D_1 \setminus D_2, w \in D_2 \setminus D_1$ – and there are two distinct edges – ab, cd – in $E(G)$, such that the following properties are satisfied: 1) ab and cd are private edges of u_1 and u_2 (with respect to D_1), respectively; 2) w ve -dominates both edges – ab and cd .

Claim 8. $u_1 \notin \{c, d\}, u_2 \notin \{a, b\}$ and $u_1 \nprec \{c, d\}, u_2 \nprec \{a, b\}$.

Proof. This is obvious from the definition of a private edge.

In all claims below we use the fact that $u_1 \nprec u_2$, since D_1 is an independent set.

Claim 9. The edges ab and cd are not adjacent in G .

Proof. Assume, to the contrary, that ab and cd are adjacent. Without loss of generality, let us suppose that $b=c$. By claim 8, $u_1 \nprec \{c, d\}$ and $u_2 \nprec \{a, b\}$. Hence, $b \nprec \{u_1, u_2\}$. Since u_1 ve -dominates ab and u_2 ve -dominates cd , we have that $u_1 \sim a$ and $u_2 \sim d$. If $a \nprec d$, then $\{a, b, d, u_1\}$ induces G_1 , a contradiction. Thus, $a \sim d$. But then $\{a, d, u_1, u_2\}$ induces G_1 . This contradiction completes the proof of the claim.

Claim 10. The vertex w is not incident to any edge in $\{ab, cd\}$.

Proof. Assume, to the contrary, that w is incident to some edge in $\{ab, cd\}$. By claim 9, edges ab and cd are not adjacent in G . Hence, taking into account the symmetry, we may assume without loss of generality that $w=b$. By symmetry, there are four cases depending on whether or not vertices u_1 and u_2 belong to $\{a, b\}$ and $\{c, d\}$, respectively.

Case 1. $u_1 = a, u_2 = c$. Since w ve -dominates cd , we have $b \sim d$. Taking into account claim 8, we conclude that $a \nprec d$ and $b \nprec c$. Thus, $\{a, b, c, d\}$ induces G_1 , a contradiction.

Case 2. $u_1 = a, u_2 \notin \{c, d\}$. By claim 8, $a \nprec \{c, d\}$ and $u_2 \nprec \{a, b\}$. Since u_2 ve -dominates cd , we may assume without loss of generality that $u_2 \sim c$. Then $b \nprec c$, otherwise $\{a, b, c, u_2\}$ induces G_1 , which is a contradiction. Since w ve -dominates cd , we have that $b \sim d$. But then $\{a, b, c, d\}$ induces G_1 , a contradiction.

Case 3. $u_1 \notin \{a, b\}, u_2 = c$. By claim 8, $u_1 \nprec \{c, d\}$ and $c \nprec \{a, b\}$. Since u_1 ve -dominates ab , we may assume without loss of generality that $u_1 \sim a$. Since w ve -dominates cd , we have $b \sim d$. Then $a \sim d$, otherwise $\{a, b, c, d\}$ induces G_1 , which is a contradiction. But now $\{a, c, d, u_1\}$ induces G_1 , a contradiction.

Case 4. $u_1 \notin \{a, b\}, u_2 \notin \{c, d\}$. By claim 8, $u_1 \nprec \{c, d\}$ and $u_2 \nprec \{a, b\}$. Since u_2 ve -dominates cd , u_2 must have a neighbour in the set $\{c, d\}$. Due to symmetry, we may assume without loss of generality that $u_2 \sim c$. Since u_1 ve -dominates ab , u_1 must have a neighbour in the set $\{a, b\}$.

Suppose that $u_1 \sim a$. Then $a \nprec c$, otherwise $\{a, c, u_1, u_2\}$ induces G_1 , which is a contradiction. Since w ve -dominates cd , w must have a neighbour in the set $\{c, d\}$. If $w \sim c$, then $\{a, b, c, u_2\}$ induces G_1 , a contradiction. Hence, $w \nprec c$ and, as consequence, $w \sim d$. But then G contains either induced $G_1 = G(a, b, c, d)$ (if $a \nprec d$) or induced $G_1 = G(a, c, d, u_1)$ (if $a \sim d$), a contradiction.

Now let us suppose that $u_1 \nprec a$. Hence, $u_1 \sim b$. But then G contains either induced $G_1 = G(b, c, u_1, u_2)$ (if $w \sim c$) or induced $G_1 = G(b, c, d, u_1)$ (if $w \nprec c, w \sim d$), a contradiction. Claim 10 is proved.

Claim 11. The vertices u_1 and u_2 are not incident to edges ab and cd , respectively.

Proof. By claim 9, edges ab and cd are not adjacent in G , and by claim 10, $w \notin \{a, b, c, d\}$. Assume, to the contrary, that u_1 is incident to ab or u_2 is incident to cd . Due to symmetry, there are two cases depending on whether or not the vertices u_1 and u_2 belong to $\{a, b\}$ and $\{c, d\}$, respectively.

Case 1. $u_1 = a, u_2 = c$. In this case $b \nprec d$, otherwise $\{a, b, c, d\}$ induces G_1 a contradiction. Since w ve -dominates both edges – ab and cd , w has at least one neighbour in each of $\{a, b\}$ and $\{c, d\}$.

Suppose that $w \nprec \{b, d\}$. Then $w \sim \{a, c\}$ and, as consequence, $G_1 = G(a, b, c, w)$ and $G_1 = G(a, c, d, w)$, a contradiction. Hence, w has a neighbour in the set $\{b, d\}$. Due to symmetry, we may assume, without loss of generality, that $w \sim b$. Let $w \nprec d$. Then clearly $w \sim c$ and, as consequence, $\{b, c, d, w\}$ induces G_1 , a contradiction.

Thus, $w \sim \{b, d\}$. Next, we conclude that $w \sim \{a, c\}$, since otherwise $G_1 = G(a, b, d, w)$ or $G_1 = G(b, c, d, w)$, which again gives a contradiction. But now, the set $\{a, b, c, d, w\}$ induces G_2 , a contradiction.

Case 2. $u_1 = a, u_2 \notin \{c, d\}$. Since u_2 *ve*-dominates cd , u_2 must have a neighbour in the set $\{c, d\}$. Due to symmetry, we may assume without loss of generality that $u_2 \sim c$. In this case $b \not\sim c$, otherwise $\{a, b, c, u_2\}$ induces G_1 , a contradiction. If $b \sim d$, then $\{a, b, c, d\}$ induces G_1 , a contradiction. Hence, $b \not\sim d$. By considering the subgraph $H = G(a, b, c, d, w)$, one can prove that H contains either G_1 or G_2 as an induced subgraph. The proof is similar to that of case 1 of this claim and hence omitted. Claim 11 is proved.

We now consider the general case, in which a, b, c, d, u_1, u_2, w are seven distinct vertices. Since u_1 *ve*-dominates ab , we may assume without loss of generality that $u_1 \sim a$. Due to symmetry, we may assume without loss of generality that $u_2 \sim c$ (since u_2 *ve*-dominates cd). By claim 8, $u_1 \not\sim \{c, d\}$ and $u_2 \not\sim \{a, b\}$. If $a \sim c$, then $\{a, c, u_1, u_2\}$ induces G_1 , a contradiction. Hence, $a \not\sim c$. Since w *ve*-dominates both edges – ab and cd , w has at least one neighbour in each of $\{a, b\}$ and $\{c, d\}$.

If $w \not\sim \{b, d\}$, then $w \sim \{a, c\}$. In this case we conclude that $w \sim \{u_1, u_2\}$, since otherwise $G_1 = G(a, c, u_1, w)$ or $G_1 = G(a, c, u_2, w)$, which gives a contradiction. But now, the set $\{a, c, u_1, u_2, w\}$ induces G_2 , a contradiction. Hence, w has a neighbour in the set $\{b, d\}$. Due to symmetry, we may assume, without loss of generality, that $w \sim b$. If $w \sim c$, then $w \not\sim a$, otherwise we have the previous situation when $w \sim \{a, c\}$. But then G contains either induced $G_1 = G(a, b, c, u_2)$ (if $b \sim c$) or induced $G_1 = G(a, b, c, w)$ (if $b \not\sim c$), a contradiction.

Thus, we have $w \sim \{b, d\}$ and $w \not\sim \{a, c\}$. If $b \sim c$ then $\{a, b, c, u_2\}$ induces G_1 , a contradiction. Hence, $b \not\sim c$. Similarly we have $a \not\sim d$, otherwise $\{a, c, d, u_1\}$ induces G_1 , a contradiction. Finally, we conclude that G contains either induced $G_1 = G(a, b, c, d)$ (if $b \sim d$) or induced $G_1 = G(a, b, d, w)$ (if $b \not\sim d$). This contradiction completes the proof of the claim.

Thus, we have contradictions in all possible cases, and the proof of the implication (ii) $\Rightarrow$ (i) is complete. Theorem 6 is proved.

By theorem 6, $W_{ve}(1)$ contains the class of $\{G_1, G_2\}$ -free graphs. On the other hand, any hereditary subclass of $W_{ve}(1)$ is contained in the class of $\{G_1, G_2\}$ -free graphs, since graphs G_1 and G_2 do not belong to $W_{ve}(1)$. Thus, we obtain the following corollary.

Corollary 1. *The unique maximal hereditary subclass of $W_{ve}(1)$ is the class of $\{G_1, G_2\}$ -free graphs.*

Finally, we consider a hereditary subclass of the class $W_{ev}(1)$. A graph G is called perfectly well *ev*-covered if every isolate-free induced subgraph of G is well *ev*-covered. The following result can be obtained using a similar technique as in the proof of theorem 6.

Theorem 7. *For a graph G , the following statements are equivalent:*

- (i) *G is a perfectly well *ev*-covered graph;*
- (ii) *G is a $\{G_1, G_2, G_3\}$ -free graph, where the graphs G_1, G_2 and G_3 are shown in fig. 3.*

By theorem 7, $W_{ev}(1)$ contains the class of $\{G_1, G_2, G_3\}$ -free graphs. On the other hand, any hereditary subclass of $W_{ev}(1)$ is contained in the class of $\{G_1, G_2, G_3\}$ -free graphs, since graphs G_1, G_2 and G_3 do not belong to $W_{ev}(1)$. Thus, we obtain the following corollary.

Corollary 2. *The unique maximal hereditary subclass of $W_{ev}(1)$ is the class of $\{G_1, G_2, G_3\}$ -free graphs.*

As a simple consequence of theorems 6 and 7, we have the following corollary.

Corollary 3. *The classes of perfectly well *ve*-covered graphs and perfectly well *ev*-covered graphs are recognisable in polynomial time.*

Related results

Consider the following decision problem.

Independent *ve*-Dominating Set problem

Instance: a graph G and an integer k .

Question: is $i_{ve}(G) \leq k$?

The Independent *ve*-Dominating Set problem is known to be NP-complete for general graphs. Moreover, it remains NP-complete even for graphs having a specific structure, such as bipartite graphs and chordal graphs [17]. On the other hand, the problem can be solved in polynomial time for trees [17].

A graph G is said to be 2- ve -indominated if there exists an integer $k \geq 1$ such that $|D| \in \{k, k+1\}$ for every minimal independent ve -dominating set D in G . The proof of theorem 1 implies the following result.

Theorem 8. *The Independent ve -Dominated Set problem is an NP-complete problem for 2- ve -indominated $K_{1,4}$ -free graphs.*

Proof. The problem of deciding « $i_{ve}(G) \leq k$ » belongs to NP. Let $I = (X, Y, Z, M)$ be an instance of the 3-DM problem. Set $k = |X| = |Y| = |Z|$. For the $K_{1,4}$ -free graph $G(I)$ in the proof of theorem 1, we have $i_{ve}(G(I)) \geq k$ and $\alpha_{ve}(G(I)) = k+1$. Hence, $G(I)$ is a 2- ve -indominated graph. Moreover, the instance $I = (X, Y, Z, M)$ is solvable for the 3-DM problem if and only if $i_{ve}(G(I)) = k$. Theorem 8 is proved.

Now consider the following decision problem.

Independent ev -Dominated Set problem

Instance: a graph G and an integer k .

Question: is $i_{ev}(G) \leq k$?

The Independent ev -Dominated Set problem is known to be NP-complete for general graphs. Moreover, it remains NP-complete for bipartite graphs [17]. On the other hand, the problem can be solved in polynomial time for trees [17].

A graph G is said to be 2- ev -indominated if there exists an integer $k \geq 1$ such that $|E'| \in \{k, k+1\}$ for every minimal independent ev -dominating set E' in G . The proof of theorem 3 implies the following result.

Theorem 9. *The Independent ev -Dominated Set problem is an NP-complete problem for 2- ev -indominated graphs in the class of graphs $A \cap B$.*

Proof. The problem of deciding « $i_{ev}(G) \leq k$ » belongs to NP. Let $J = (C, X)$ be an instance of the 3Sat problem. Set $k = n = |X|$. For the graph $G(J) \in A \cap B$ in the proof of theorem 3, we have $i_{ev}(G(J)) \geq k$ and $\alpha_{ev}(G(J)) = k+1$. Hence, $G(J)$ is a 2- ev -indominated graph. Moreover, there exists a satisfying truth assignment for C if and only if $i_{ev}(G(J)) = k$. Theorem 9 is proved.

Since $\gamma_{ev}(G) = i_{ev}(G)$ for any isolate-free graph G (see [17]), we can use theorem 9 to show that the problem of deciding « $\gamma_{ev}(G) \leq k$ » is NP-complete in the class of graphs $A \cap B$.

Consider the concept of paired-domination in graphs, which was introduced by T. Haynes and P. Slater [26; 27] (see [28] for a survey).

A subset $D \subseteq V(G)$ is a paired-dominating set of G if every vertex of $V(G)$ is adjacent to some vertex in D and the induced subgraph $G(D)$ contains a perfect matchings (not necessarily induced). Note that every graph without isolated vertices has a paired-dominating set. It is easy to see that for any paired-dominating set D of G , the cardinality $|D|$ is even. A paired-dominating set D in G is minimal if no proper subset of D is a paired-dominating set. The paired domination number of G , denoted by $\gamma_{pr}(G)$, is the minimum cardinality of a paired-dominating set of G . The upper paired domination number of G , denoted by $\Gamma_{pr}(G)$ is the maximum cardinality of a minimal paired-dominating set of G .

The following characterisation of a minimal paired-dominating set is obtained in work [29].

Proposition 1. *Let G be an isolate-free graph and D be a paired-dominating set of G . Then D is a minimal paired-dominating set if and only if any two vertices $x, y \in D$ satisfy one of the following conditions:*

- (i) *the induced subgraph $G(D \setminus \{x, y\})$ does not contain a perfect matching;*
- (ii) *without loss of generality, x is an end-vertex in $G(D)$ adjacent to y ;*
- (iii) *there exists a vertex $u \in V(G) \setminus D$ such that $N(u) \cap D \subseteq \{x, y\}$.*

The following result can be found in paper [23].

Proposition 2. *If G is an isolate-free graph, then $2i_{ev}(G) = \gamma_{pr}(G) \leq \Gamma_{pr}(G) \leq 2\alpha_{ev}(G)$.*

We call a graph well paired-dominated if all its minimal paired-dominating sets are of the same cardinality (cf. [30] for the difference in definitions). That is, all minimal paired-dominating sets are also minimum paired-dominating sets. In other words, a graph G is well paired-dominated, if and only if $\gamma_{pr}(G) = \Gamma_{pr}(G)$ is satisfied.

Consider the following decision problem.

Non-Well Paired-Dominated Graph problem

Instance: a graph G .

Question: is it true that the graph G is not well paired-dominated? In other words, are there two minimal paired-dominating sets of distinct cardinalities in G ?

Using theorem 3, one can show that the Non-Well Paired-Dominated Graph problem is an NP-complete problem.

Theorem 10. *The Non-Well Paired-Dominated Graph problem is NP-complete in the class of graphs $A \cap B$.*

Proof. It can be checked in polynomial time whether a paired-dominating set is minimal. This follows from the characterisation of a minimal paired-dominating set given in proposition 1 and the fact that the general perfect matching problem can be solved in polynomial time (see, for example, [31]). Thus, the Non-Well Paired-Dominated Graph problem is in NP since we can guess subsets $D_1, D_2 \subseteq V(G)$ and, in polynomial time, verify that D_1 and D_2 are minimal paired-dominating sets of distinct cardinalities. To show NP-completeness, we use exactly the same arguments as in the proof of theorem 3.

Given an instance $J = (C, X)$ of the 3Sat problem, construct the graph $G = G(J) \in A \cap B$ as in the proof of theorem 3. Let D be a minimal paired-dominating set of G . Since the cardinality $|D|$ is even, $i_{ev}(G) \geq n$, $\alpha_{ev}(G) = n + 1$ and $\gamma_{pr}(G) \leq |D| \leq \Gamma_{pr}(G)$, by using proposition 2 we obtain that $|D| \in \{2n, 2n + 2\}$. The graph G has always a minimal paired-dominating set of cardinality $2n + 2$, for example $\{c_1, c_2\} \cup \{y_i, z_i \mid i = 1, 2, \dots, n\}$ is such a set. Moreover, there exists a satisfying truth assignment for C if and only if $i_{ev}(G) = n$. But since $\gamma_{pr}(G) = 2i_{ev}(G) = 2n$, we also obtain that there exists a satisfying truth assignment for C if and only if $\gamma_{pr}(G) = 2n$, i. e. if and only if the graph G has a minimal paired-dominating set of cardinality $2n$. Theorem 10 is proved.

Conclusions

In this paper, we prove that the problem of recognising well ve -covered graphs is co-NP-complete. We show that the co-NP-completeness result holds for $K_{1,4}$ -free graphs as well. Also, we prove that the problem of recognising well ev -covered graphs is still co-NP-complete for some classes defined by certain forbidden induced subgraphs such as $\{C_4, \overline{K_1 \cup C_4}, K_{1,5}, 2P_4\}$ -free graphs. We introduce new hereditary classes of graphs, the perfectly well ve -covered graphs and the perfectly well ev -covered graphs, and we characterise them by finite sets of forbidden induced subgraphs, thus obtaining new polynomial-time recognisable classes, where parameters $i_{ve}(G)$, $\alpha_{ve}(G)$ and $i_{ev}(G)$, $\alpha_{ev}(G)$ are easy to compute. We prove that the problem of deciding « $i_{ve}(G) \leq k$ » (« $i_{ev}(G) \leq k$ », respectively) is NP-complete even if the graph has minimal independent ve -dominating (minimal independent ev -dominating, respectively) sets of at most two cardinalities differing by one. Finally, we introduce a new class of graphs, the well paired-dominated graphs, and we show that recognising a well paired-dominated graph is a co-NP-complete problem.

An interesting open problem is to investigate the computational complexity of the Non-Well ve -Covered Graph problem (Non-Well ev -Covered Graph problem, respectively) for $K_{1,3}$ -free graphs ($K_{1,t}$ -free graphs, where $t \in \{3, 4\}$, respectively). Note that our approach does not apply to this case.

References

1. Bondy JA, Murty USR. *Graph theory*. Berlin: Springer; 2008. XII, 651 p. (Graduate texts in mathematics; volume 244).
2. Garey MR, Johnson DS. *Computers and intractability: a guide to the theory of NP-completeness*. San Francisco: W. H. Freeman and Company; 1979. XII, 338 p.
3. Haynes T, Hedetniemi ST, Henning MA. *Domination in graphs: core concepts*. Cham: Springer; 2023. XX, 644 p. (Springer monographs in mathematics). DOI: 10.1007/978-3-031-09496-5.
4. Plummer MD. Some covering concepts in graphs. *Journal of Combinatorial Theory*. 1970;8(1):91–98. DOI: 10.1016/S0021-9800(70)80011-4.
5. Chvátal V, Slater PJ. A note on well-covered graphs. *Annals of Discrete Mathematics*. 1993;55:179–181. DOI: 10.1016/S0167-5060(08)70387-X.
6. Sankaranarayana RS, Stewart LK. Complexity results for well-covered graphs. *Networks*. 1992;22(3):247–262. DOI: 10.1002/net.3230220304.
7. Caro Y, Sebö A, Tarsi M. Recognizing greedy structures. *Journal of Algorithms*. 1996;20(1):137–156. DOI: 10.1006/jagm.1996.0006.
8. Brown J, Hoshino R. Well-covered circulant graphs. *Discrete Mathematics*. 2011;311(4):244–251. DOI: 10.1016/j.disc.2010.11.007.
9. Baptiste P, Kovalyov MY, Orlovich YuL, Werner F, Zverovich IE. Graphs with maximal induced matchings of the same size. *Discrete Applied Mathematics*. 2017;216(part 1):15–28. DOI: 10.1016/j.dam.2016.08.015.
10. Tankus D, Tarsi M. Well-covered claw-free graphs. *Journal of Combinatorial Theory. Series B*. 1996;66(2):293–302. DOI: 10.1006/jctb.1996.0022.
11. Tankus D, Tarsi M. The structure of well-covered graphs and the complexity of their recognition problems. *Journal of Combinatorial Theory. Series B*. 1997;69(2):230–233. DOI: 10.1006/jctb.1996.1742.
12. Finbow A, Hartnell B, Nowakowski RJ. A characterization of well-covered graphs of girth 5 or greater. *Journal of Combinatorial Theory. Series B*. 1993;57(1):44–68. DOI: 10.1006/jctb.1993.1005.

13. Finbow A, Hartnell B, Nowakowski RJ. A characterization of well-covered graphs that contain neither 4- nor 5-cycles. *Journal of Graph Theory*. 1994;18(7):713–721. DOI: 10.1002/jgt.3190180707.
14. Caro Y, Ellingham MN, Ramey JE. Local structure when all maximal independent sets have equal weight. *SIAM Journal on Discrete Mathematics*. 1998;11(4):644–654. DOI: 10.1137/S0895480196300479.
15. Prisner E, Topp J, Vestergaard PD. Well-covered simplicial, chordal, and circular arc graphs. *Journal of Graph Theory*. 1996;21(2):113–119. DOI: 10.1002/(SICI)1097-0118(199602)21:2<113::AID-JGT1>3.0.CO;2-U.
16. Peters KW. *Theoretical and algorithmic results on domination and connectivity* [dissertation]. Clemson: Clemson University; 1986. IX, 162 p.
17. Lewis JR. *Vertex-edge and edge-vertex domination in graphs* [dissertation]. Clemson: Clemson University; 2007. XVI, 82 p.
18. Lewis JR, Hedetniemi ST, Haynes TW, Fricke GH. Vertex-edge domination. *Utilitas Mathematica*. 2010;81:193–213.
19. Jamieson AC. Linear algorithms for edge-vertex domination in trees. *Congressus Numerantium*. 2007;187:214–222.
20. Boutrig R, Chellali M, Haynes TW, Hedetniemi ST. Vertex-edge domination in graphs. *Aequationes mathematicae*. 2016;90(2):355–366. DOI: 10.1007/s00010-015-0354-2.
21. Boutrig R, Chellali M, Meddah N. Well ve-covered graphs. *Communications in Combinatorics and Optimization*. 2025;10(1):69–78.
22. Boutrig R, Chellali M. Well ev-covered trees. *RAIRO – Operations Research*. 2023;57(3):1481–1489. DOI: 10.1051/ro/2023088.
23. Chellali M. Further results on independent edge-vertex domination. *Discrete Mathematics, Algorithms and Applications*. 2024;16(6):2350080. DOI: 10.1142/S1793830923500805.
24. Kitaev S, Lozin V. *Words and graphs*. Cham: Springer; 2015. XVIII, 264 p. (Monographs in theoretical computer science). DOI: 10.1007/978-3-319-25859-1.
25. Dabrowski K, Lozin VV, Zamaraev V. On factorial properties of chordal bipartite graphs. *Discrete Mathematics*. 2012;312(16):2457–2465. DOI: 10.1016/j.disc.2012.04.010.
26. Haynes TW, Slater PJ. Paired-domination and the paired-domatic number. *Congressus Numerantium*. 1995;109:65–72.
27. Haynes TW, Slater PJ. Paired-domination in graphs. *Networks*. 1998;32(3):199–206. DOI: 10.1002/(SICI)1097-0037(199810)32:3<199::AID-NET4>3.0.CO;2-F.
28. Desormeaux WJ, Haynes TW, Henning MA. Paired domination in graphs. In: Haynes TW, Hedetniemi ST, Henning MA, editors. *Topics in domination in graphs*. Cham: Springer; 2020. p. 31–77 (Developments in mathematics; volume 64). DOI: 10.1007/978-3-030-51117-3_3.
29. Edwards M. *Criticality concepts for paired domination in graphs* [dissertation]. Victoria: University of Victoria; 2006. VIII, 75 p.
30. Fitzpatrick SL, Hartnell BL. Well paired-dominated graphs. *Journal of Combinatorial Optimization*. 2010;20(2):194–204. DOI: 10.1007/s10878-008-9203-8.
31. Lawler E. *Combinatorial optimization: networks and matroids*. Mineola: Dover Publications; 2001. X, 374 p.

Received 29.01.2026 / revised 17.02.2026 / accepted 24.02.2026.

УДК 004.932

СИСТЕМА ДЕТЕКТИРОВАНИЯ ДЫМА И САБОТАЖА НА БАЗЕ ОДНОПЛАТНОГО КОМПЬЮТЕРА ДЛЯ МОНИТОРИНГА ПОЖАРОВ В РЕЖИМЕ РЕАЛЬНОГО ВРЕМЕНИ

Е. Р. АДАМОВСКИЙ¹⁾, Р. П. БОГУШ¹⁾

¹⁾Полоцкий государственный университет им. Евфросинии Полоцкой,
ул. Блохина, 29, 211440, г. Новополоцк, Беларусь

Аннотация. Представлена архитектура системы видеонаблюдения с использованием периферийных вычислений для обнаружения дыма и саботажа. Система состоит из параллельно функционирующих модулей, отвечающих за управление конфигурацией и обработку видеоданных, включая их прием, анализ и передачу. Детектирование дыма осуществляется с помощью алгоритмов компьютерного зрения и базируется на последовательном извлечении его визуальных признаков: направления движения, цветовых характеристик и влияния на фоновые объекты. Саботаж обнаруживается путем оценки резкости сцены с применением фильтра для выделения границ объектов. Обработка видеопотока высокого разрешения выполняется с использованием одноплатного компьютера, при этом производится распределение задач по ядрам на многоядерном процессоре для повышения производительности. Такое решение существенно снижает нагрузку на центральный узел системы, особенно при масштабировании на большое число камер. Результаты экспериментов демонстрируют высокую точность обнаружения дыма и саботажа, а также стабильную работу архитектуры в режиме реального времени.

Ключевые слова: детектирование дыма; компьютерное зрение; системы видеонаблюдения; одноплатный компьютер.

Образец цитирования:

Адамовский ЕР, Богуш РП. Система детектирования дыма и саботажа на базе одноплатного компьютера для мониторинга пожаров в режиме реального времени. *Журнал Белорусского государственного университета. Математика. Информатика*. 2026;1:108–120 (на англ.).
EDN: WTXHNV

For citation:

Adamovskiy YR, Bohush RP. Smoke and camera tampering detection system based on single-board computer for real-time fire monitoring. *Journal of the Belarusian State University. Mathematics and Informatics*. 2026;1:108–120.
EDN: WTXHNV

Авторы:

Егор Русланович Адамовский – старший преподаватель кафедры вычислительных систем и сетей факультета информационных технологий.
Рихард Петрович Богуш – доктор технических наук, профессор; заведующий кафедрой вычислительных систем и сетей факультета информационных технологий.

Authors:

Yahor R. Adamovskiy, senior lecturer at the department of computing systems and networks, faculty of information technologies.
e.adamovsky@psu.by
<https://orcid.org/0000-0003-1044-8741>
Rykhard P. Bohush, doctor of science (engineering), full professor; head of the department of computing systems and networks, faculty of information technologies.
bogushr@mail.ru
<https://orcid.org/0000-0002-6609-5810>

SMOKE AND CAMERA TAMPERING DETECTION SYSTEM BASED ON SINGLE-BOARD COMPUTER FOR REAL-TIME FIRE MONITORING

Y. R. ADAMOVSKIY^a, R. P. BOHUSH^a

^a*Euphrosyne Polotskaya State University of Polotsk,
29 Blahina Street, Navapolack 211440, Belarus*

Corresponding author: Y. R. Adamovskiy (e.adamovsky@psu.by)

Abstract. An edge-based architecture for a video surveillance system with integrated smoke and camera tampering detection is presented. The system consists of parallel-executing modules responsible for configuration management and video data processing, including receive, analysis, and transmission. Smoke detection is performed using computer vision algorithms based on the sequential extraction of discriminative visual features: motion direction, colour characteristics, and impact on background visibility. Camera tampering is detected by evaluating scene sharpness through an edge-detection filter. High-resolution video processing is performed on a single-board computer, with computational tasks distributed across the cores of a multi-core central processor to optimise performance. The proposed approach significantly reduces the load on the central node of the surveillance infrastructure, which is a critical advantage for scalable deployments involving numerous cameras. Experimental results demonstrate high detection accuracy for smoke and tampering, as well as stable real-time operation of the implemented architecture.

Keywords: smoke detection; computer vision; video surveillance systems; single-board computer.

Introduction

Fire detection is a critical task with a wide range of applications, including residential and industrial facilities, streets, and forested areas. In its early stage, a fire can be easily mitigated; however, its further development and spread may lead to significant damage. Therefore, monitoring the initial phases of a fire is of considerable interest. Computer vision algorithms that analyse video data represent a promising approach to early fire detection, including in outdoor environments. Smoke is often the dominant or even the only visible indicator of fire at an early stage [1]. It propagates faster than flames and occupies a comparatively larger volume of space because it does not require close contact with the high-temperature region of material combustion. Therefore, when developing an early-stage fire detection algorithm based on video frame analysis, it is reasonable to focus on smoke detection.

Visually, smoke is characterised by the following features [2]: vertical upward motion of particles that transitions into chaotic motion; a limited colour range in shades of gray; semi-transparency, which causes blurring of background objects; a quasi-stable source location; an irregular shape with flickering. These features can be extracted by analysing video frames obtained from a surveillance system in real time. However, video data is subject to various distortions and limitations, including limited information due to camera resolution, colour range, and frame rate; loss of scene depth; sensor noise; optical system effects; variable lighting conditions; compression artifacts introduced by lossy video codecs. Therefore, when developing a smoke detection algorithm, it is essential to exploit as many of these features as possible through joint spatial and temporal processing to achieve high detection accuracy (AC) and minimise errors.

When video surveillance systems with smoke detection capabilities are deployed at industrial facilities, it becomes necessary to rapidly identify anomalies that cause the video signal to lack relevant information about the monitored scene or about the signal itself, which may indicate tampering. Tampering refers to an intentional action aimed at disrupting the operation of a video surveillance system while preserving the appearance of normal functionality. Tampering may involve physical interference with cameras, for example by covering their lenses with opaque or translucent objects, painting over them, or overexposing the image sensor, which renders it impossible to capture intruder activity or detect visual smoke features. Algorithms for this task must account for complex conditions encountered in real-world operation [3]: the appearance of new objects in the scene that occupy a significant portion of the frame; camera motion induced by wind; variations in lighting, time of day, and weather conditions.

Video surveillance systems are typically large-scale infrastructures equipped with numerous cameras. Continuous monitoring of the proper operation of a large number of cameras over extended periods can be challenging for security personnel and operators. Consequently, there is a risk of significant material damage due to either a missed fire in its early stage or undetected tampering. The large number of surveillance devices results in insufficient computing resources for centralised real-time processing of the resulting data volumes. One possible solution is edge computing, which offloads computation to edge devices that operate in conjunction with video cameras and perform initial analysis and processing of the video stream.

Related works

The concept of an edge or embedded system for threat detection, including fire and camera tampering, involves processing the video stream directly on the camera or endpoint device, as such devices are often deployed in remote or hard-to-access locations. Therefore, the computing unit must be compact and energy-efficient, which inherently limits its computational capacity. Existing implementations typically employ single-board computers (SBC), such as the Raspberry Pi and NVIDIA Jetson, as the hardware platform.

Fire and camera tampering can be regarded as a complex threat detectable by computer vision techniques in video surveillance systems. Certain steps in smoke and camera tampering detection algorithms are similar, as both aim to extract informative features from video frames to determine the presence of a threat.

Smoke detection methods deployed in embedded systems are either adapted to operate on low-performance hardware or replaced with novel algorithms designed to minimise computational complexity. For instance, in work [4] the authors evaluate the use of artificial neural networks (ANN) for edge-based fire detection across several architectures, including VGG19, ResNet18, EfficientNetV2, Vision Transformer (ViT), and ConvNeXt, on an NVIDIA Jetson Xavier platform. Experimental results demonstrate that model compression techniques, such as quantisation, pruning, and knowledge distillation, enable these networks to achieve a frame-level classification AC of 95 % while maintaining processing rates above 30 fps. Nevertheless, the Jetson Xavier significantly outperforms and exceeds the cost of typical SBC, suggesting that deploying high-complexity ANN in resource-constrained edge environments remains impractical.

In article [5] the compact MobileNetV2 model was employed for smoke and flame detection in forest fire scenarios. The model operated on a Raspberry Pi 5 (further – RPi5), processing an input video stream at 30 fps. It was trained on a dataset comprising 40 000 images and optimised through quantisation, input cropping, and conversion to TensorFlow Lite format. The system achieved a processing speed of 1–2 fps, with an F1-score of 94 % on a held-out test set. To suppress false positives, the system generated a class prediction for the current frame based on the five preceding frames and transmitted the result to the server.

In work [6] a portable smoke detection system based on the Raspberry Pi 4B (further – RPi4B) was developed. It comprised a Raspberry Pi camera, a video frame processing module, a detector, and display and notification subsystems. The data processing module included a 5×5 Gaussian filter for noise reduction and histogram equalisation for contrast enhancement. Smoke detection was performed using a cascade AdaBoost classifier based on local binary pattern (LBP) features, extracted from each video frame via a sliding window approach. The classifier was trained on a dataset of 6800 image patches of size 35×35 and 48×48 px using an NVIDIA GTX 1080 graphics processing unit. The implementation was carried out in Python using the OpenCV computer vision library. The system processes video frames of 320×240 px at 25 fps, achieving an AC of 94.7 %. The reduced input resolution, resulting from LBP feature extraction, was imposed by hardware constraints. Output data was saved to files and accessed remotely, while smoke detection alerts were delivered via email. It should be noted that real-time operation in research [6] relied on low-resolution frames, which limits the effective detection of small objects. Moreover, temporal information that could have contributed to improved detection AC was not exploited.

In article [7] a multi-stage smoke detection algorithm is proposed, which includes the following steps: motion detection using a Kalman filter; saturation segmentation in the HSV colour space; combination of the resulting binary masks using the logical AND operator; identification of regions of interest (RoI) with a sufficiently large area and a specified aspect ratio range; classification of the regions using a lightweight convolutional neural network (CNN) with an input dimension of $50 \times 50 \times 3$ and a detection threshold of 0.7. The model was trained on a dataset of 14 000 images. To suppress false positives, smoke was considered detected only if its presence was confirmed continuously for at least 3 s. The implementation was developed in Python using the Keras machine learning framework. The algorithm was evaluated on NVIDIA Jetson Nano (JN) and Raspberry Pi 3 (further – RPi3) platforms, achieving an average AC of 87.7 % at processing speeds of 55 and 9 fps, respectively, with an input frame resolution of 320×240 px. A high false positive rate was observed, which may result from suboptimal alarm triggering duration or limited generalisation capability of the employed CNN. Furthermore, the algorithm's performance on the RPi3 is insufficient for real-time operation. Key features of the approach include incorporation of temporal context and preprocessing aimed at reducing data volume and computational load on the CNN.

A CNN was also employed for fire detection based on the YOLOv2 architecture and implemented on JN [8]. The implementation was developed in MATLAB using CUDA technology. The model was trained on a dataset of 400 images of size 128×128 px and achieved an F1-score of 95.4 % on test videos at a processing speed of 21 fps. Fires were detected 1–2 s after their actual occurrence. It should be noted that video frames were downsampled to extremely low resolution, yet real-time processing was not achieved. The algorithm incurs

significant computational costs and relies on proprietary CUDA technology developed by NVIDIA, which restricts its applicability to other SBC.

The YOLOv8 architecture was employed for fire detection with several adaptations for the Ingenic T41 (further – T41) low-power embedded system-on-chip: reparameterisation, replacement of the SiLU activation function with LeakyReLU to simplify model quantisation, pretraining on the COCO dataset, fine-tuning on the domain-specific D-Fire dataset comprising 21 500 images, and minimisation of data merging and splitting operations [9]. The trained model achieved 79.4 % mAP at a processing speed of 47 fps. Despite these strong results, it should be noted that the optimisations were tailored to a specific hardware architecture, and their effectiveness may diminish when deployed on other SBC.

Another approach using YOLO for forest fire detection with unmanned aerial vehicles is presented in work [10]. The YOLOv5 architecture was deployed on the RPi4B with the following modifications: replacement of the CNN in the backbone with a lightweight ShuffleNetV2 model, network thinning, and subsequent retraining. The model was trained on a dataset of 24 400 images of size 320×240 px using an NVIDIA RTX 3050 graphics processing unit. The CPU clock rate was increased by 33 %. The system achieved a processing speed of 9 fps with 92.5 % mAP. Comparison with [9] shows that the modifications in [10] were less platform-specific, yet delivered lower processing speed despite using smaller input resolution, even though the RPi4B offers superior computing power compared to the T41. This indicates that hardware-aware optimisations can significantly enhance algorithm performance in computer vision tasks.

An analysis of the YOLO and CNN-RCNN architectures for UAV-based fire detection on the JN was conducted in [11]. The YOLOv8n model was employed for object detection, while a hybrid CNN-RCNN model performed classification using temporal information. Training was carried out on a dataset of 2947 images of size 640×640 px. The system achieved a processing speed of 12–20 fps with an F1-score of 88 %. The results indicate that the use of a sequence of multiple ANN did not yield high performance; however, the approach demonstrated good AC in smoke detection.

In work [12] tampering detection is performed by evaluating the sharpness of object boundaries in images using a Canny edge detector. The results are analysed over time, and a Kalman filter is applied to suppress false alarms. The following tampering scenarios were considered: defocusing, spray-painting, occlusion, and camera redirection. The precision (P) of the proposed method was 96.6 %.

Video camera tampering detection is performed using deep neural networks in article [13]. The authors employ the AlexNet, ResNet18, ResNet50, and DenseNet161 architectures. AlexNet achieved the highest AC at 74.9 %, but generated over 2000 false alarms per hour, which is unacceptable for real-world deployment. The authors note that incorporating temporal context would improve the results. Moreover, the computational requirements of the considered neural networks preclude their real-time operation on SBC.

The analysis shows that existing algorithmic solutions for smoke and camera tampering detection are generally unsuitable for deployment on SBC, when high-resolution video must be processed in real time. Therefore, the smoke detection algorithms and models employed on these devices reduce the input data size, which improves performance but may lead to missed detections of small objects. The solutions discussed employ lightweight, trainable classifiers and modified ANN, the combination of which with classical computer vision algorithms and hardware-specific optimisations can improve the performance of smoke detection algorithms. To reduce false positives, some approaches accumulate extracted features over several seconds and base smoke detection decisions on the aggregated data. However, most methods based on ANN rely solely on information from the current video frame, thereby neglecting temporal context. Both rule-based and neural network-based methods can be employed for camera tampering detection. However, the use of ANN is limited by their high computational complexity. The limited computing power of SBC restricts the number of services that can run simultaneously, thereby preventing complex video analysis in addition to smoke detection. Therefore, developing an effective smoke and camera tampering detection system with high performance and low computational requirements for embedded video surveillance modules remains a pressing challenge.

Efficient real-time smoke and tampering detection video system

System architecture. The proposed edge-based system for smoke and camera tampering detection comprises server and client components. The server component includes a SBC, specifically a RPi4B, an IP camera, and a network switch. The client component consists of an operator workstation equipped with dedicated software, including video surveillance tools and a web application for interaction with the server. The block diagram of the system is shown in fig. 1.

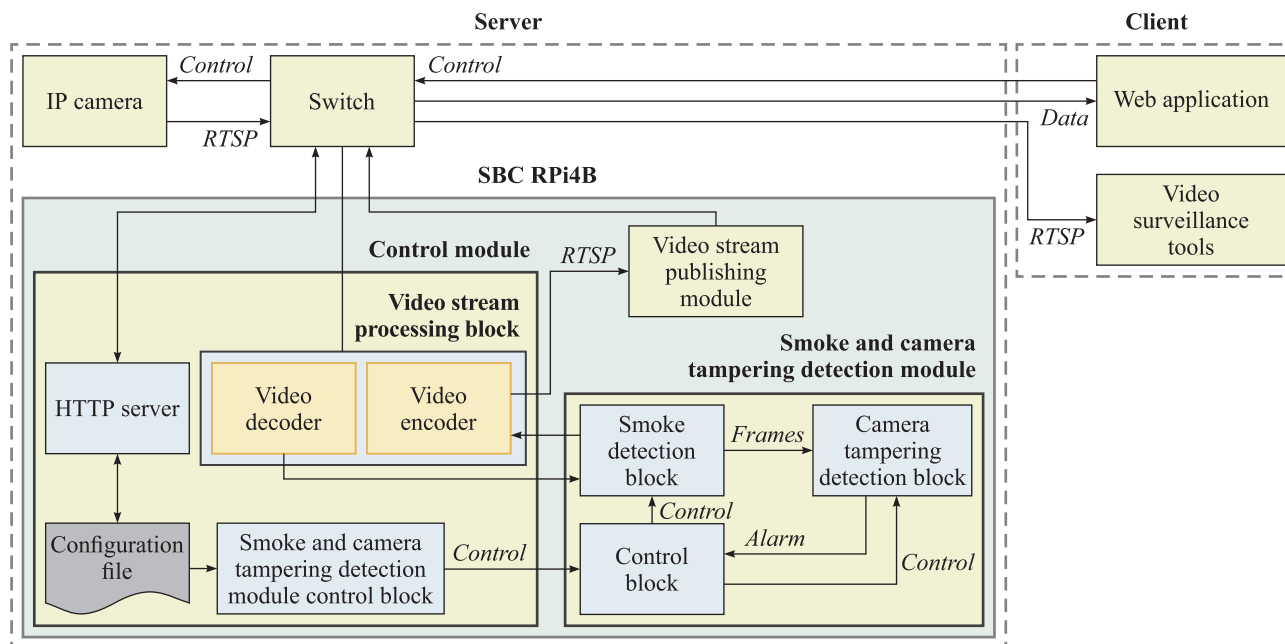


Fig. 1. Edge computing architecture for real-time smoke and camera tampering detection on a RPi4B

The server-side software stack deployed on the SBC comprises three core modules: a control module, a smoke and camera tampering detection module, and a video stream publishing module.

The functionality of the control module can be divided into interconnected blocks: a client web application control block, which is essentially a hypertext transfer protocol (HTTP) server, a smoke and camera tampering detection module control block, and a video stream processing block. The control module loads parameters from the configuration file, including general settings, operational parameters for the smoke and camera tampering detection module, and the real-time streaming protocol uniform resource locator of the video stream source. The web application control block processes HTTP requests and provides file server functionality. These requests include updating the configuration file, serving static web content, transmitting the current video frame, and delivering archived frames.

The smoke and camera tampering detection module control block launches the detectors, applies the configuration settings, and monitors their operation. The video stream processing block decodes the input video stream and forwards frames to the smoke and camera tampering detection module. It then receives the processed frames from this module, encodes them, and transmits them to the video stream publishing module. The smoke and camera tampering detection module comprises three blocks: a smoke detection block, a camera tampering detection block, and a control block that manages data exchange between the two.

The processed video stream publishing module makes the frames available for viewing on the operator workstation using standard video surveillance tools.

System operation algorithm. The proposed system operates using the developed algorithms for smoke and camera tampering detection. The smoke detection algorithm sequentially applies computer vision methods to the original video frames to extract discriminative smoke features. The processing order is determined by two factors: computational complexity and the discriminative power of the feature for smoke. If no features are detected at a given stage, subsequent stages are skipped, thereby reducing the computational load on the device. A key aspect of the approach is the division of video frames into a regular grid, with features averaged within each cell. This strategy simplifies and accelerates computation through uniform processing and parallelisation.

The camera tampering detection block, designed for real-time video processing on the SBC, exploits a characteristic property of video surveillance cameras: their minimum focusing distance is typically tens of centimetres or more, depending on the lens. Therefore, if the camera is obscured by a nearby object, such as a sheet of paper, image sharpness is reduced, even if the sheet reproduces the original scene. This reduction in sharpness can be detected automatically.

The system operation algorithm describes the modules and functional blocks, including the smoke and camera tampering detection techniques, and consists of the following steps.

1. The control module is launched. The module loads parameters from the configuration file. The HTTP server is enabled to support the client web application, and initiates periodic broadcast messages on the local network to allow the operator to discover the RPi4B and access the web application.

2. The smoke and camera tampering detection module is launched. Its control block loads the parameters from the configuration file obtained in step 1. In the event of an emergency shutdown of the detectors, the module automatically restarts. A restart is also triggered if the configuration file is updated via the HTTP server.

3. The video stream processing block receives an input video stream from a source, whose uniform resource locator is specified in the configuration file, decodes it, and transmits each frame I_t to the smoke and camera tampering detection module. If the input video stream becomes unavailable, the block generates a synthetic video stream containing an error message and notifies the web application control block.

4. The smoke and camera tampering detection module receives the frame I_t as input and launches detectors that process it according to the described algorithms.

4.1. Operating algorithm of the smoke detection block [14] within the smoke and camera tampering detection module.

4.1.1. The module is initialised in standby mode upon system startup.

4.1.2. Each frame I_t is resised to a fixed spatial dimension for further processing.

4.1.3. The histogram of I_t is normalised.

4.1.4. An initial background model B is constructed using the MOG2 algorithm [15].

4.1.5. The module transitions to operating mode after a background initialisation period T_b , required to construct the initial background model.

4.1.6. Cells C_m containing motion are extracted using background subtraction followed by morphological erosion and dilation as post-processing steps.

4.1.7. The background model B is updated, excluding cells where motion was detected in frame I_{t-1} . If smoke is detected, the update is suspended entirely.

4.1.8. Cells $C_{m.st}$ exhibiting stable motion over a time window T_m are selected.

4.1.9. Cells C_c , whose colour characteristics correspond to those of smoke, are selected.

4.1.9.1. Frame I_t is converted to the HSV colour space, yielding I_t^{HSV} .

4.1.9.2. Threshold binarisation is applied to I_t^{HSV} , S and V channels separately. The colour properties of smoke are more effectively captured in the HSV colour space, enabling robust segmentation [16].

4.1.9.3. Masks obtained in step 4.1.9.2 are combined using the logical AND operator to produce the final smoke candidate mask C_c .

4.1.10. Candidate cells C_1 of the primary RoI are computed by combining the results of steps 4.1.8 and 4.1.9 into a weight array. If a cell in the current iteration exhibits steady motion and matches smoke in colour, its weight in the array is set to the maximum value. Otherwise, its weight is decremented by a fixed step in each subsequent iteration until it reaches zero. A candidate cell is defined as any cell in the weight array with a non-zero value. Further processing is performed only on the I_t fragment that encompasses all current candidate cells within a bounding rectangle.

4.1.11. The selection of cells C_{hf} , for which a significant drop in the energy of high-frequency components relative to the background model B , is detected. For this purpose a two-dimensional wavelet transform is applied [17].

4.1.12. Cells $C_{hf.st}$ exhibiting a stable reduction in high-frequency energy over a time window T_{hf} are selected.

4.1.13. Selection of cells C_v , for which a sufficient contrast reduction relative to the background model, is detected.

4.1.14. Candidate cells C_2 for the secondary RoI are computed by applying the logical AND operator to C_1 and the results of the previous steps 4.1.12 and 4.1.13. Further processing is performed only for the fragment I_t that encompasses all candidate cells within a bounding rectangle.

4.1.15. Cells C_f exhibiting bottom-up motion are selected using the optical flow algorithm [18] applied to C_2 regions in frames I_t and I_{t-1} . A cell is selected if its motion vector lies within a specified angular range and exceeds a predefined amplitude threshold.

4.1.16. Cells $C_{f.st}$ exhibiting stable bottom-up motion over a time window T_f are retained.

4.1.17. Sets $C_{f.st}$ and C_2 are combined using the logical AND operator to produce C_r .

4.1.18. Cells C_a are selected by accumulating C_r over time and comparing the accumulated value with a presence decision threshold.

4.1.19. A decision on smoke presence is made if at least one non-zero element exists in C_a .

4.2. Operating algorithm of the camera tampering detection block: periodic analysis of video frames for tampering indicators during normal operation.

4.2.1. Frame I_t is converted to a grayscale image I_t^{bw} .

4.2.2. The Laplacian L_t of I_t^{bw} is computed by convolving the image with the Laplacian kernel. Each element of L_t represents the second-order spatial derivative at the corresponding pixel $I_t^{bw}(x, y)$.

4.2.3. The standard deviation σ for L_t is computed and used as a metric for image sharpness assessment.

4.2.4. Camera tampering in the frame is declared if $\sigma < \sigma_{th}$ is met, where σ_{th} is a predefined threshold.

4.2.5. The number of consecutive frames N with detected tampering is counted.

- 4.2.6. Camera tampering is confirmed if N exceeds the threshold N_{th} .
4.2.7. If camera tampering is not detected, then steps 4.2.1–4.2.6 are repeated, otherwise step 4.2.8 is executed.
4.2.8. The module enters tampering mode: smoke detection is suspended, and the output video stream is annotated with a tampering alert.
4.2.9. Steps 4.2.1–4.2.6 are repeated until camera tampering is no longer detected.
4.2.10. The control block halts the module. Subsequently, the control block of the smoke and camera tampering detection module restarts it, as described in step 2.

Multi-threaded real-time video processing architecture on quad-core embedded platform. The control module is implemented in the Rust programming language. The smoke and camera tampering detection module is implemented in C++ using the OpenCV library. The stream publishing module is implemented using the MediaMTX media server.

To improve performance, data processing steps are executed in parallel across the four CPU cores of the RPi4B with tasks distributed among dedicated threads to ensure efficient operation and real-time processing.

Thus, video stream transmission, encoding, and decoding are performed on CPU cores 0 and 1. On CPU core 2, thread 0 receives decoded frames in the smoke and camera tampering detection module and executes the tampering detection. Thread 1 is dedicated exclusively to motion detection using the MOG2 algorithm, as this operation is computationally intensive. On CPU core 3, thread 0 executes the remaining stages of the smoke detection algorithm, and thread 1 transmits processed frames back to the video stream processing module.

The block diagram of the proposed model implementation, reflecting the distribution of tasks across CPU cores, is shown in fig. 2.

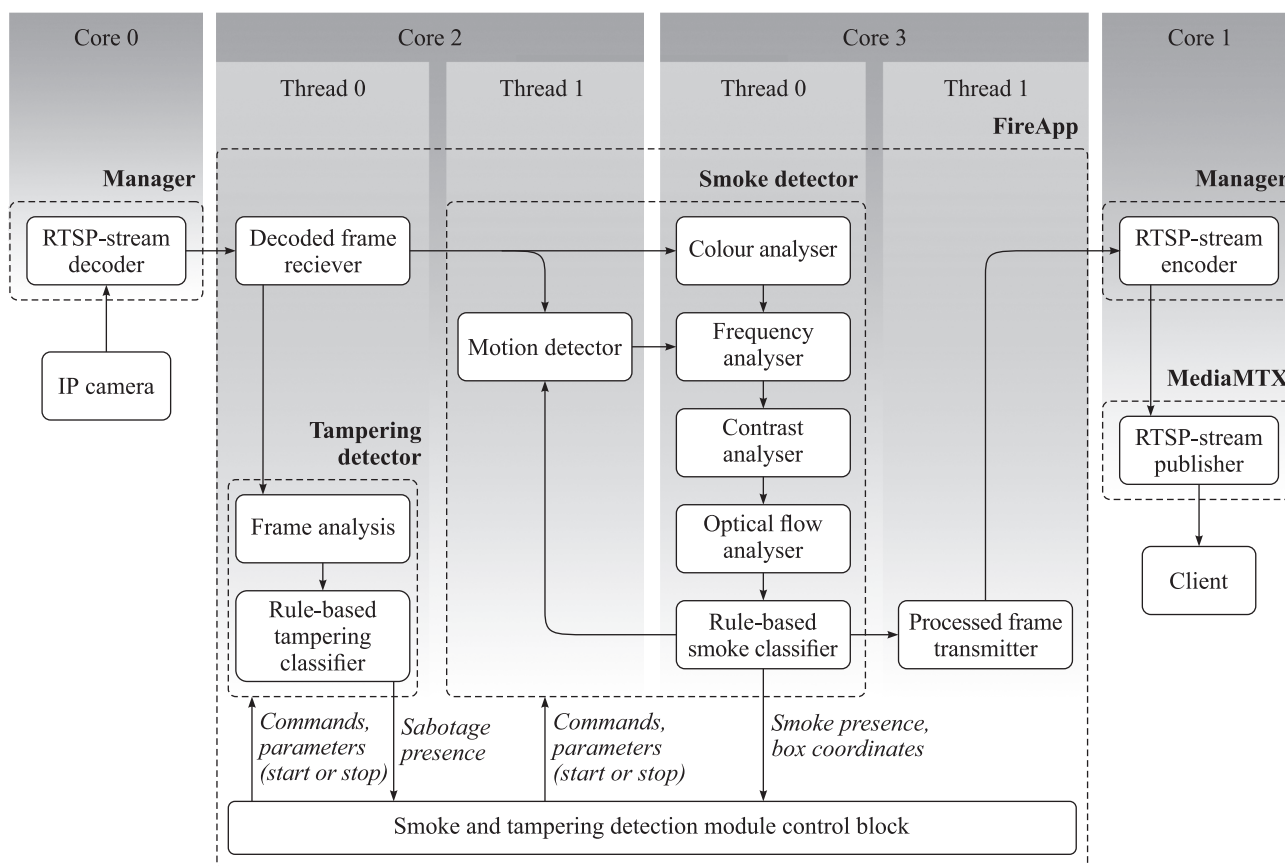


Fig. 2. Block diagram of the distribution of tasks in a model based on a 4-core CPU SBC

Experimental results

Performance evaluation of the smoke detection algorithm. To evaluate the performance of the algorithm, a dataset comprising video recordings of smoke and smoke-like phenomena at resolutions up to 1280×720 px (HD) was collected from open sources. Additional recordings were captured by the authors to supplement the dataset.

Experimental results show that the proposed smoke detection algorithm achieves an F1-score of 97.9 % and processes HD video streams at 25 fps. The use of spatiotemporal analysis and a combination of visual smoke features reduces the error rate to an average of 2 %, which is lower than that of comparable methods [2].

Table 1 presents the evaluation criteria for smoke detection methods, and the last row presents the results of testing the proposed algorithm. The compared metrics include AC, P, false-positive rate (FPR), false-negative rate (FNR) and F1-score [1]. Additionally, the table specifies the processed frame size, processing performance, and the SBC used.

Table 1

Comparison of smoke detection algorithms in edge-based systems

Method	Hard-ware	Frame size, px	FPS	Metrics, %					Source
				AC	FPR	FNR	F1-score	mAP	
MobileNetV2	RPi5	224 × 224	1–2	97.9	–	–	98.3	–	[5]
AdaBoost+	RPi4B	320 × 240	25	94.7	1.1	3.6	–	–	[6]
Rule-based+	RPi3	320 × 240	9	87.7	24.7	4.7	90.6	–	[7]
	JN	320 × 240	55						
YOLOv2	JN	128 × 128	21	96.8	3.4	2.9	95.4	–	[8]
YOLOv8n+	JN	640 × 640	20	89.0	6.5	2	88.0	–	[11]
YOLOv8	T41	672 × 416	47	–	–	–	–	79.4	[9]
YOLOv5	RPi4B	320 × 240	9	–	–	–	–	92.5	[10]
Proposed algorithm	RPi4B	1280 × 720	25	97.9	0.26	1.77	97.5	–	This work

Figure 3 shows examples of detection on video frames from a prepared dataset, where detected smoke regions are indicated by bounding boxes.



Fig. 3. Smoke detection results on video frames produced by the proposed algorithm

Performance evaluation of the camera tampering detection. The camera tampering detection module was evaluated using a custom dataset comprising video sequences of four scene types: daytime street, nighttime street, illuminated indoor environment, and semi-dark indoor environment (fig. 4).

The proposed method was compared with two other approaches based on image entropy estimation: the Shannon entropy calculation algorithm, which uses the distribution of pixel intensities represented as a histogram, and entropy estimation based on the compression ratio of the Lempel – Ziv – Markov chain algorithm (further – LZMA). The LZMA is deterministic, has efficient implementations in programming languages and is used in the 7z archive format. The LZMA compression ratio, defined as the ratio of the original data size to the compressed size, is inversely proportional to the entropy of the observed scene. Therefore, it provides a convenient and reproducible means for entropy estimation [19].

Four camera tampering scenarios were tested: covering the camera lens with a light-coloured object, covering the camera lens with a dark object, covering the camera lens with a translucent object, and illuminating the lens with a bright light source (fig. 5).



Fig. 4. Scene types for tampering detection evaluation:
a – daytime street; *b* – nighttime street;
c – illuminated indoor scene; *d* – semi-dark indoor scene



Fig. 5. Tampering scenarios for detector evaluation:
a – covering the camera lens with a light-coloured object;
b – covering the camera lens with a dark object;
c – covering the camera lens with a translucent object;
d – illuminating the lens with a bright light source

The obtained results are presented in table 2. The ranges of estimated values for individual test video recordings for different types of sabotage are shown.

Table 2

Estimated performance of camera tampering detection algorithms

Criterion	Scene (see fig. 4)	Reference estimated value	Estimated value in tampering scenarios (see fig. 5)			
			<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>
Sharpness assessment (σ)	<i>a</i>	39.20–57.70	22.50–24.70	19.90–24.20	17.10–19.20	–
	<i>b</i>	19.10–19.90	13.80–15.10	–	11.10–12.70	12.90–16.20
	<i>c</i>	25.0–26.6	16.20–17.60	15.70–16.60	13.0–14.0	17.20–20.50
	<i>d</i>	20.80–21.90	12.9–16.0	13.50–14.70	9.09–10.2	8.73–9.89
Shannon entropy estimate (H), bit/px	<i>a</i>	7.01–7.09	6.56–6.73	6.69–6.86	6.78–6.83	–
	<i>b</i>	5.91–5.93	3.96–4.28	–	6.50–6.55	7.52–7.71
	<i>c</i>	7.10–7.11	6.04–6.12	6.74–6.81	6.84–6.94	6.39–6.50
	<i>d</i>	5.51–5.54	5.57–5.66	2.74–3.32	5.79–5.97	6.80–7.26
LZMA-based entropy estimation (k)	<i>a</i>	2.46–2.57	4.68–4.94	5.02–5.57	8.28–8.85	–
	<i>b</i>	4.31–4.41	11.90–13.50	–	6.01–6.19	3.73–6.72
	<i>c</i>	4.67–4.83	8.41–9.12	5.28–5.74	8.38–9.94	5.51–8.31
	<i>d</i>	10.7–12.0	8.55–9.55	16.40–19.40	9.63–11.10	8.05–10.60

The reference estimated value in table 2 was obtained from video frames without features of camera tampering.

An analysis of table 2 shows that, for certain scenes, the outputs of algorithms 2 and 3 can be identical in the presence and absence of tampering. In contrast, the proposed tampering detection method, which is based on frame sharpness, enables the determination of a scene-specific threshold value σ_{th} for unambiguous classification of tampering presence or absence.

Thermal stability and real-time performance evaluation. The third and equally important stage of the research is to determine whether a SBC can operate within its thermal limits while maintaining real-time performance. The evaluation was performed under three load conditions of the smoke detection block.

1. No load (mode 1): the observed scene contains no smoke-like objects. Only the initial algorithm stages, namely motion detection and colour-based object detection, are active.

2. Partial load (mode 2): the observed scene contains smoke-like objects. Smoke regions are either absent or occupy a small portion of the frame for brief durations. All algorithm stages are periodically active, but detection events are infrequent and short-lived.

3. Full load (mode 3): the observed scene contains smoke that occupies a significant portion of the frame over an extended duration, representing the maximum computational load on all algorithm stages.

Video sequences corresponding to the described load conditions were selected from the dataset, and testing demonstrated that the system maintains a processing rate of 25 fps during extended operation.

The CPU temperature was recorded at 1-second intervals under ambient room temperature conditions, using a device housed in a case with a passive heatsink and no active cooling. The results were smoothed using a 60-sample moving average window to reduce noise and are presented in fig. 6.

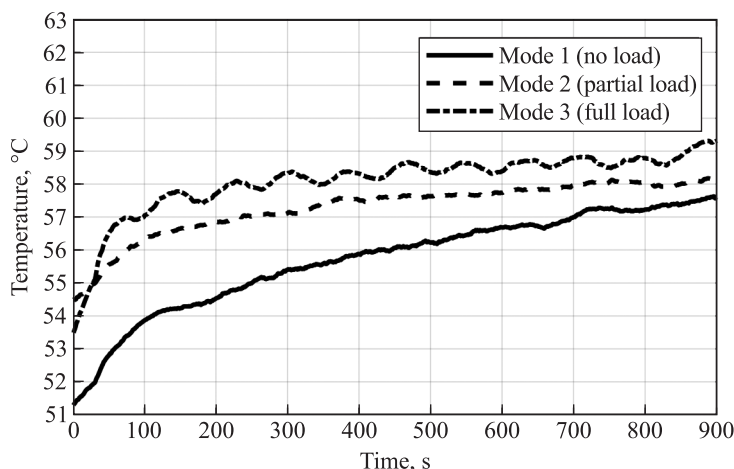


Fig. 6. RPi4B CPU temperature under various load conditions

Figure 6 shows that the initial CPU temperature is 50 °C. The peak temperatures observed during testing were 58.9 °C in mode 1, 59.4 °C in mode 2, and 60.4 °C in mode 3. The maximum permissible CPU temperature for the RPi4B is 80 °C, which is not exceeded in any load mode.

CPU core utilisation was also measured across all load modes with tampering detection disabled. The measurement data is presented in fig. 7 as histograms with core utilisation grouped into 10 % intervals. The height of each bar represents the number of measurements falling within the corresponding utilisation interval.

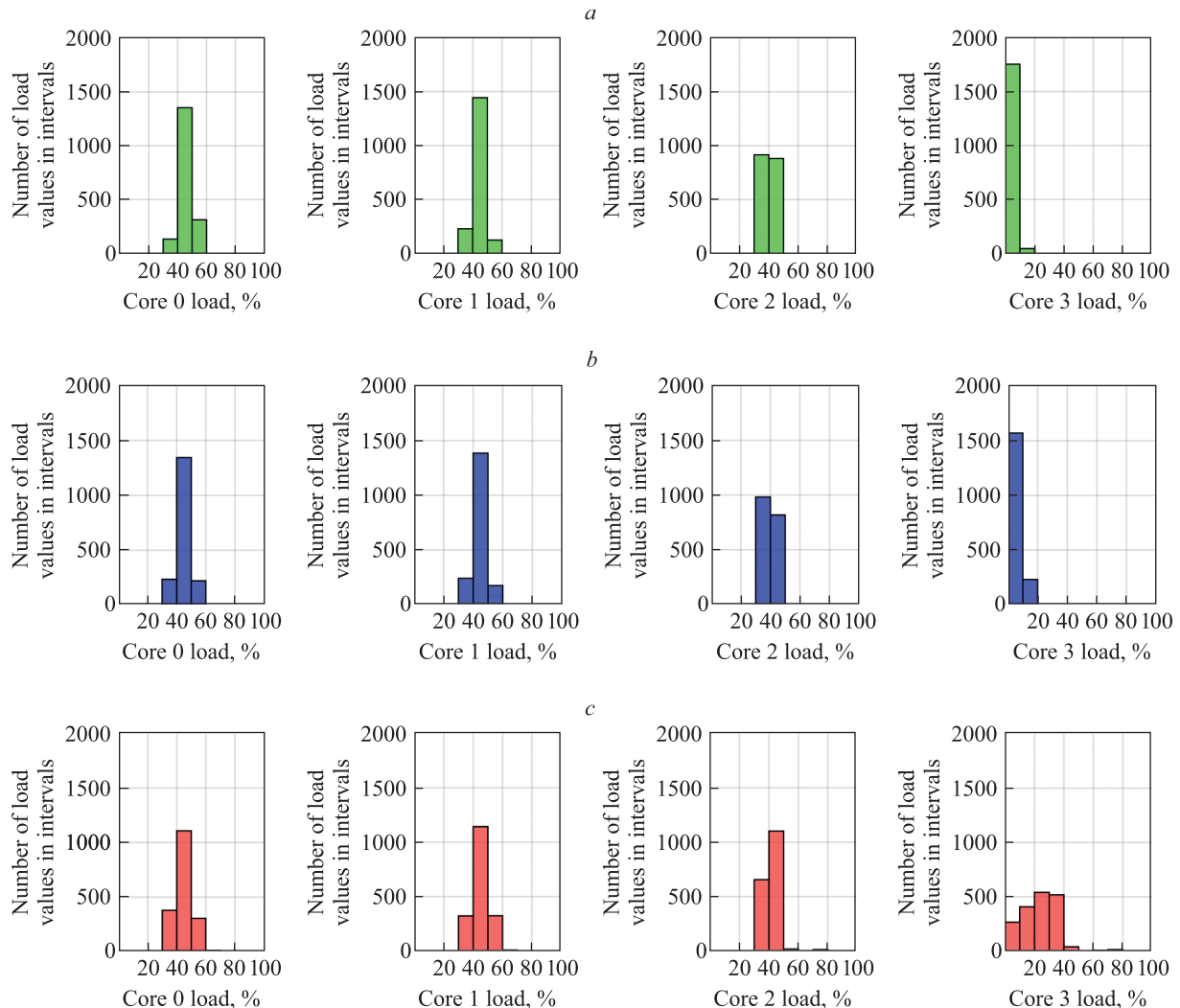


Fig. 7. RPi4B CPU core load: *a* – mode 1; *b* – mode 2; *c* – mode 3

A comparison of the data in fig. 7 shows that CPU cores 0 and 1 are consistently utilised at 30–60 %, regardless of the load mode, as they are not employed by the smoke detection algorithm. Core 2 is utilised at an average of 30–50 % and executes tampering and motion detection, which is the initial stage of the smoke detection algorithm and is applied to every frame. The utilisation of core 3 is the most mode-dependent, as it executes all subsequent stages of the smoke detection algorithm. Its load increases from 10–20 % in mode 1 to 40 % in mode 3. None of the cores reaches 100 % load, indicating a balanced distribution of tasks across the cores and ensuring stable system operation.

To assess the impact of the camera tampering detector on system performance, the CPU utilisation of core 2 was measured with the detector enabled and disabled. The results show that even in mode 3 (full load) core utilisation ranges from 31.2 to 48.6 %, which matches the range shown in fig. 7, *c*. Therefore, the tampering detector does not adversely affect the stability of the smoke detection algorithm or other system modules.

The experimental results demonstrate the effectiveness of the proposed smoke and camera tampering detection architecture. Analysis of the results confirms that both the algorithms and the overall system based on the proposed architecture, implemented on a SBC, perform robustly across a range of operating conditions.

Conclusions

This paper presents an edge-based architecture for a video surveillance system equipped with smoke and camera tampering detection capabilities, enabling real-time processing and analysis of high-resolution video streams on a RPi4B single-board computer. The system can be seamlessly integrated into existing video surveillance networks. By leveraging decentralised computing, the proposed approach significantly reduces the computational load on the central node, which is essential for scalable deployments involving a large number of cameras.

Experimental results confirm correct system operation, high detection AC for both smoke and tampering, and robust performance during continuous operation. The smoke detection algorithm achieves an AC of 97.9 % while processing high-quality video at 25 fps. The use of spatiotemporal analysis of smoke-related visual features reduces the total error rate to an average of 2.03 %, outperforming comparable state-of-the-art methods. Camera tampering detection based on image sharpness assessment via edge-based standard deviation enables reliable threshold selection without degrading the processing speed of the smoke detector or the overall system performance. Testing of the complete system demonstrated that the developed algorithms do not disrupt the thermal balance of the single-board computer, enabling deployment of the system in industrial applications.

References

1. Chaturvedi Sh, Khanna P, Ojha A. A survey on vision-based outdoor smoke detection techniques for environmental safety. *ISPRS Journal of Photogrammetry and Remote Sensing*. 2022;185:158–187. DOI: 10.1016/j.isprsjprs.2022.01.013.
2. Adamovskiy Y, Bohush R. Real-time algorithm for light gray smoke detection in video sequences. In: Shmaliy YS, editor. *8th International conference on computing, control and industrial engineering; 2024 June 21–22; Wuhan, China. Volume 1*. Singapore: Springer; 2024. p. 535–542 (Lecture notes in electrical engineering; volume 1252). DOI: 10.1007/978-981-97-6934-6_64.
3. Vennam P, Pramod TC, Thippeswamy BM, Kim Y-G, Kumar PBN. Attacks and preventive measures on video surveillance systems: a review. *Applied Sciences*. 2021;11(12):5571. DOI: 10.3390/app11125571.
4. Mahmoudi SA, Gloesener M, Benkedadra M, Lerat J-S. Edge AI system for real-time and explainable forest fire detection using compressed deep learning models. In: Bashford-Rogers T, Meneveaux D, Ammi M, Ziat M, Jänicke S, Purchase H, et al., editors. *Proceedings of the 20th International joint conference on computer vision, imaging and computer graphics theory and applications; 2025 February 26–28; Porto, Portugal. Volume 3*. Setubal: SciTePress; 2025. p. 847–854. DOI: 10.5220/0013382500003912.
5. Sharobiddinov D, Siddiqui HUR, Saleem AA, Mendez Mezquita G, Ramirez Vargas DL, de la Torre Díez I. Edge-based autonomous fire and smoke detection using MobileNetV2. *Sensors*. 2025;25(20):6419. DOI: 10.3390/s25206419.
6. Liu B, Sun B, Cheng P, Huang Y. An embedded portable lightweight platform for real-time early smoke detection. *Sensors*. 2022;22(12):4655. DOI: 10.3390/s22124655.
7. Gagliardi A, de Gioia F, Saponara S. A real-time video smoke detection algorithm based on Kalman filter and CNN. *Journal of Real-Time Image Processing*. 2021;18(6):2085–2095. DOI: 10.1007/s11554-021-01094-y.
8. Saponara S, Elhanashi A, Gagliardi A. Real-time video fire/smoke detection based on CNN in antifire surveillance systems. *Journal of Real-Time Image Processing*. 2021;18(3):889–900. DOI: 10.1007/s11554-020-01044-0.
9. Peng R, Cui C, Wu Y. Real-time fire detection algorithm on low-power endpoint device. *Journal of Real-Time Image Processing*. 2025;22(1):29. DOI: 10.1007/s11554-024-01605-7.
10. Ye J, Ioannou S, Nikolaou P, Raspopoulos M. CNN based real-time forest fire detection system for low-power embedded devices. In: Institute of Electrical and Electronics Engineers. *Proceedings of the 31st Mediterranean conference on control and automation; 2023 June 26–28; Limassol, Cyprus*. Piscataway: Institute of Electrical and Electronics Engineers; 2023. p. 137–143. DOI: 10.1109/MED59994.2023.10185692.
11. Shamta I, Demir BE. Development of a deep learning-based surveillance system for forest fire detection and monitoring using UAV. *PLoS ONE*. 2024;19(3):0299058. DOI: 10.1371/journal.pone.0299058.
12. Wang Y-K, Fan C-T, Cheng YK, Deng PS. Real-time camera anomaly detection for real-world video surveillance. In: Institute of Electrical and Electronics Engineers. *Proceedings of the 2011 International conference on machine learning and cybernetics; 2011 July 10–13; Guilin, China. Volume 4*. Piscataway: Institute of Electrical and Electronics Engineers; 2011. p. 1520–1525. DOI: 10.1109/ICMLC.2011.6017032.
13. Mantini P, Shah SK. UHCTD: a comprehensive dataset for camera tampering detection. In: Institute of Electrical and Electronics Engineers. *Proceedings of the 16th IEEE International conference on advanced video and signal based surveillance; 2019 September 18–21; Taipei, Taiwan*. Piscataway: Institute of Electrical and Electronics Engineers; 2019. p. 1–8. DOI: 10.1109/AVSS.2019.8909856.
14. Ye S, Adamovskiy Y, Chen H, Bohush R, Ablameyko S. Real-time smoke detection in video based on two-step selection of regions of interest and directional movement analysis. *Pattern Recognition and Image Analysis*. 2025;34:1323–1332. DOI: 10.1134/S1054661824701396.
15. Naimeng C, Wanjuan Y, Xiaoyu W. Smoke detection for early forest fire in aerial photography based on GMM background and wavelet energy. In: Institute of Electrical and Electronics Engineers. *Proceedings of the 2021 IEEE International conference on cower electronics, computer applications; 2021 January 22–24; Shenyang, China*. Piscataway: Institute of Electrical and Electronics Engineers; 2021. p. 763–765. DOI: 10.1109/ICPECA51329.2021.9362647.

16. Fan X, Lei F, Yang K. Real-time detection of smoke and fire in the wild using unmanned aerial vehicle remote sensing imagery. *Forests*. 2025;16(2):201. DOI: 10.3390/f16020201.
17. Ye S, Bai Z, Chen H, Bohush R, Ablameyko S. An effective algorithm to detect both smoke and flame using color and wavelet analysis. *Pattern Recognition and Image Analysis*. 2017;27(1):131–138. DOI: 10.1134/S1054661817010138.
18. Wu Y, Chen M, Wo Y, Han G. Video smoke detection based on dense optical flow and convolutional neural network. *Multimedia Tools and Applications*. 2021;80(6):35887–35901. DOI: 10.1007/s11042-020-09870-x.
19. Challapalli K, Kothamasu V, Garikapati H, Charan RS, Thota M, Anamalamudi S. Analysing lossless image compression techniques for IoT devices. In: Institute of Electrical and Electronics Engineers. *Proceedings of the 2024 Third International conference on electrical, electronics, information and communication technologies; 2024 July 24–26; Trichirappalli, India*. Piscataway: Institute of Electrical and Electronics Engineers; 2024. p. 1–6. DOI: 10.1109/ICEEICT61591.2024.10718516.

Received 24.12.2025 / revised 19.01.2026 / accepted 29.01.2026.

АНАЛИЗ УПРАВЛЕНЧЕСКИХ РЕШЕНИЙ
С ИСПОЛЬЗОВАНИЕМ МАЛЫХ ЯЗЫКОВЫХ МОДЕЛЕЙК. В. АНДРЕНКО¹⁾, А. А. КРОЩЕНКО¹⁾, В. А. ГОЛОВКО¹⁾, Ю. Б. КРАПИВИН¹⁾¹⁾Брестский государственный технический университет,
ул. Московская, 267, 224017, г. Брест, Беларусь

Аннотация. Предложен метод решения задачи автоматической оценки качества управленческих решений по текстовым ответам на русском языке, соотнесенным с проблемными ситуациями, на базе нейросетевой модели DistilBERT. Новизна исследования заключается в разработке алгоритма контекстно зависимой классификации, при котором входные данные формируются как конкатенация описания ситуации и ответа респондента, а также в использовании методики постобработки предсказаний, предполагающей взаимосвязь результатов с конкретными управленческими компетенциями или красными флагами. Достигнуты высокие показатели (ассигура = 0,986 8, F1-score = 0,987 0) при оценке управленческих решений на русском языке с интерпретируемым выводом зон профессионального и личностного роста, в то время как известные работы ориентированы на общий анализ текста или его тональности. Предложенный подход открывает возможность применения указанных алгоритма и методики при разработке систем автоматизированной оценки руководящих работников и подбора персонала, а также персонализированного развития управленческих компетенций в практике найма и корпоративном обучении.

Ключевые слова: малая языковая модель; DistilBERT; анализ управленческих решений; нейронная сеть; дистилляция знаний.

Образец цитирования:

Андренко КВ, Крощенко АА, Головки ВА, Крапивин ЮБ. Анализ управленческих решений с использованием малых языковых моделей. *Журнал Белорусского государственного университета. Математика. Информатика.* 2026;1:121–129 (на англ.).
EDN: WVIPLU

For citation:

Andrenko KV, Kroshchanka AA, Golovko VA, Krapivin YuB. Managerial decision analysis using small language models. *Journal of the Belarusian State University. Mathematics and Informatics.* 2026;1:121–129.
EDN: WVIPLU

Авторы:

Ксения Владимировна Андренко – преподаватель-стажер кафедры интеллектуальных информационных технологий факультета электронно-информационных систем.

Александр Александрович Крощенко – кандидат технических наук; доцент кафедры интеллектуальных информационных технологий факультета электронно-информационных систем.

Владимир Адамович Головки – доктор технических наук, профессор; заведующий кафедрой интеллектуальных информационных технологий факультета электронно-информационных систем.

Юрий Борисович Крапивин – кандидат технических наук; доцент кафедры интеллектуальных информационных технологий факультета электронно-информационных систем.

Authors:

Ksenia V. Andrenko, trainee lecturer at the department of intelligent information technologies, faculty of electronic and information systems.

ksenia.andrenko2002@gmail.com

<https://orcid.org/0009-0001-0976-2613>

Aliaksandr A. Kroshchanka, PhD (engineering); associate professor at the department of intelligent information technologies, faculty of electronic and information systems.

kroschenko@gmail.com

<https://orcid.org/0000-0003-3285-3545>

Vladimir A. Golovko, doctor of science (engineering), full professor; head of the department of intelligent information technologies, faculty of electronic and information systems.

vladimir.golovko@gmail.com

<https://orcid.org/0000-0003-2615-289X>

Yury B. Krapivin, PhD (engineering); associate professor at the department of intelligent information technologies, faculty of electronic and information systems.

ybox@list.ru

<https://orcid.org/0009-0004-8489-6912>



MANAGERIAL DECISION ANALYSIS USING SMALL LANGUAGE MODELS

K. V. ANDRENKO^a, A. A. KROSHCHANKA^a, V. A. GOLOVKO^a, Yu. B. KRAPIVIN^a

^aBrest State Technical University, 267 Maskowskaja Street, Brest 224017, Belarus

Corresponding author: K. V. Andrenko (ksenia.andrenko2002@gmail.com)

Abstract. This paper proposes a method for the automatic assessment of managerial decision quality based on free-form textual responses to Russian-language business cases using a fine-tuned DistilBERT model. The novelty of the study lies in a context-dependent classification algorithm with concatenated input (case description + response) and a post-processing technique that explicitly links predictions to specific managerial competencies or red flags. For the first time, high-accuracy evaluation of managerial decisions in Russian (accuracy = 0.986 8, F1-score = 0.987 0) with interpretable competency-based feedback has been achieved, unlike prior work focused on general text or sentiment analysis. The approach enables the development of automated systems for managerial assessment, personnel selection and personalised leadership development in human resources and corporate training.

Keywords: small language model; DistilBERT; management decision analysis; neural network; knowledge distillation.

Introduction

In the era of big data, the integration of artificial intelligence systems into all spheres of life, including economics, education, culture and management, has become particularly significant [1–5]. Managerial decision-making requires processing large volumes of both qualitative and quantitative information at various stages (for instance, such as during recruitment or performance appraisal to identify areas for professional growth). The sheer amount of unstructured data that cannot be handled manually has created a need for tools capable of interpreting free-form responses in the context of solving real-life managerial tasks [6]. The automatic evaluation of the quality of managerial decisions in the form of textual responses when solving case studies is therefore especially relevant.

Existing research primarily focuses on general text analysis rather than the specific assessment of managerial decision quality [7–9]. For example, the importance of applying advances in artificial intelligence and natural language processing, such as structural topic modelling, in human resources (HR) context is highlighted [10–16].

One of the most promising directions in natural language processing is the use of transformer-based models, such as BERT (*bidirectional encoder representations from transformers*) and its distilled variant, DistilBERT. The advantages of BERT are high accuracy, context awareness, and the ability to fine-tune only a single output layer to create models that solve the task of text sequence classification (for example, the classification of managerial decisions) [17–21].

DistilBERT preserves the performance of the original BERTModel while requiring substantially fewer computational resources. These features make DistilBERT a suitable model for practical applications demanding both high accuracy and low latency¹ [22; 23]. The authors previously investigated the use of distilled language models for sentiment analysis, demonstrating their effectiveness in that domain [24].

The aim of this study is to develop and evaluate a novel method for the automatic assessment of managerial decision quality based on free-form textual responses to Russian-language cases. To achieve this aim, the following tasks were addressed:

- 1) construction of a specialised Russian-language dataset, comprising 760 annotated examples of managerial cases with paired high- and low-quality responses, enriched with structured competency scores and red flags;
- 2) fine-tuning of a DistilBERT model for context-dependent binary classification by concatenating case descriptions and responses as input sequences;
- 3) comprehensive evaluation of the proposed approach using accuracy, F1-score and stratified cross-validation metrics, alongside comparisons with existing solutions.

In the present study, the DistilBERT model was fine-tuned for the binary classification of managerial decisions while explicitly accounting for case-specific context. The input sequence was formed by concatenating the case description and the respondent's textual response, thereby enabling the model to capture causal relationships and pragmatic dependencies between the situation and the proposed solution. The achieved performance metrics

¹DistilBERT // Hugging Face : website. URL: https://huggingface.co/docs/transformers/model_doc/distilbert (date of access: 19.12.2024).

(accuracy = 0.986 8, F1-score = 0.987 0) demonstrates that the proposed context-aware approach delivers excellent classification accuracy and can serve as a foundation for deployable intelligent decision-support systems in real-world managerial and HR practices.

Materials and methods

Dataset preparation. A dedicated Russian-language dataset was constructed for the binary classification of managerial decisions. It comprises 760 annotated examples corresponding to 380 unique managerial cases (Case_ID 1 – Case_ID 380). For each case, two contrasting responses were created: one high-quality response (Is_Good = 1) and one low-quality response (Is_Good = 0), yielding a perfectly balanced dataset with no class imbalance.

The responses were initially generated using three large language models (Claude 3.5 Sonnet, Grok 3 and Gemini 2.0 Pro) and subsequently underwent thorough manual revision and stylistic refinement by a professional linguist to ensure linguistic naturalness, stylistic diversity, and alignment with real-world managerial discourse.

The cases cover a broad spectrum of managerial domains typically encountered by executives at municipal and corporate levels, including infrastructure (street lighting, sidewalks, bike lanes, waste containers), social policy (support for families, elderly citizens and youth), healthcare (service conditions, medical equipment, telemedicine), education (textbooks, extracurricular activities, career guidance), public safety (vandalism, accidents, theft), ecology (waste separation, urban greening), culture and leisure (festivals, museums, parks), and transport and housing services (traffic congestion, snow removal, elevator maintenance).

Representative examples include the situation «В районе участились случаи вандализма на детских площадках» («Vandalism on children’s playgrounds has recently increased in the district») and the situation «В компании сотрудники жалуются на отсутствие гибкого графика» («Employees are complaining about the lack of flexible working hours»).

The dataset is stored in CSV format (semicolon-separated) with the following fields:

- Case_ID – unique case identifier (integer);
- Situation – textual description of the managerial problem or situation (string, mean length ≈ 78 tokens);
- Response – free-form textual response proposing a managerial decision (string);
- Is_Good – binary label (1 – good response, 0 – poor response);
- Scores – dictionary containing two competency scores (scale 3–5 points) for good responses and two competency scores (scale 1–2 points) for poor responses; stored as string representation of Python dictionary (e. g., “{ ‘Инициатива и ответственность’ : 4, ‘Планирование и организация’ : 5 }”);
- Red_Flags – dictionary containing two red flag scores: low values (1–2 points) indicating minimal shortcomings for good responses, or high values (3–5 points) indicating severe shortcomings for poor responses; stored as string representation of Python dictionary (e. g., “{ ‘Бездействие и перекладывание ответственности’ : 4, ‘Игнорирование жалоб граждан’ : 3 }”).

The inclusion of structured Scores and Red_Flags dictionaries enables not only supervised classification but also direct post-processing of model predictions into interpretable competency-based feedback and personalised development recommendations – a feature rarely found in existing text-classification datasets.

The key distinguishing feature of the dataset is the integration of structured competency scores and red flags, which extends its utility far beyond conventional binary classification datasets. This dual structure enables (i) standard supervised training, (ii) direct post-processing of model predictions into human-readable competency profiles and personalised development recommendations, and (iii) detailed interpretability analysis. These capabilities are rarely encountered in publicly available text classification corpora.

Input representation. To ensure context-aware classification, the input sequence was constructed by concatenating the case description and the respondent’s textual response, separated by the [SEP] token:

```
input_text = [CLS] + Situation + [SEP] + Response + [SEP].
```

Tokenisation was performed using the original DistilBERT tokeniser from the Hugging Face Transformers library (version 4.41.2) with the following settings: maximum sequence length = 512 tokens, truncation = True, padding = “max_length”. This approach enables the model to capture causal and pragmatic relationships between the situation and the proposed solution. For illustration, the response «I will establish a task force» may represent a good decision for the case «Residents complain about poor street lighting» but a poor decision for the case «Urgent evacuation due to fire».

Training protocol. The DistilBERT model (66 mln parameters) was fine-tuned in Google Colab using a single NVIDIA T4 GPU (VRAM – 16 GB). This hardware configuration provided sufficient memory and computational performance for the selected hyperparameters.

The dataset (760 instances) was stratified split into training (608 instances, or 80 %), validation (76 instances, or 10 %) and test (76 instances, or 10 %) sets.

All random seeds (PyTorch, NumPy, Python random) were fixed at 42 to ensure full reproducibility.

Fine-tuning was performed for three epochs using AdamW optimiser and the following hyperparameters: learning rate $2 \cdot 10^{-5}$ (linear scheduler with warmup over the first 10 % of steps), batch size 16, weight decay 0.01, gradient clipping 1.0.

Cross-entropy loss was used as the loss function, which is minimised during training to predict the `Is_Good` labels:

$$L = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\tilde{y}_i) + (1 - y_i) \log(1 - \tilde{y}_i)],$$

where N is the number of samples; y_i is the ground-truth label; $\tilde{y}_i$ is the predicted probability of the positive class.

Model architecture. The fine-tuned model consists of the pre-trained DistilBERT encoder (6 transformer layers: hidden size 768, 12 attention heads) followed by a binary classification head: a dropout layer ($p = 0.1$) and a fully connected layer with softmax activation (fig. 1).

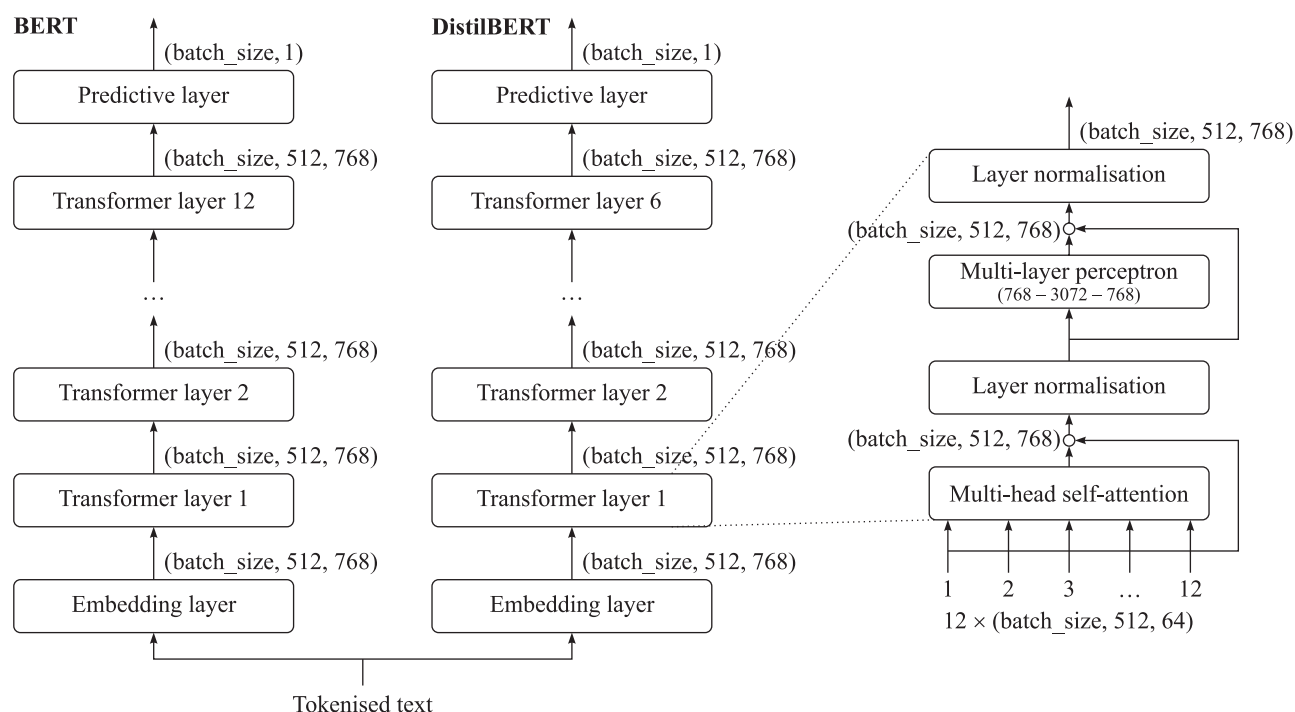


Fig. 1. Architectural comparison of BERT and DistilBERT

Each transformer layer includes the standard components: multi-head self-attention, multi-layer perceptron, layer normalisation and residual connections.

Compared with BERT-base (12 transformer layers, 110 mln parameters), DistilBERT retains approximately 96 % of the original performance while being 40 % smaller and 45 % faster in inference.

The predictive layer is a fully connected layer with one neuron and a softmax activation function, which makes the final decision on the evaluation of the decision made by the manager.

Interpretability via post-processing. After inference, predictions were enriched through deterministic post-processing:

- if `Is_Good` = 1, then the corresponding `Scores` dictionary is returned, highlighting two demonstrated competencies with high scores (3–5 points) from the universal competency framework (e. g., "{ 'Инициатива и ответственность': 4, 'Планирование и организация': 5 }");
- if `Is_Good` = 0, then the corresponding `Red_Flags` dictionary is returned, highlighting two critical shortcomings with high severity scores (3–5 points) from the red flags taxonomy (e. g., "{ 'Бездействие и перекладывание ответственности': 4, 'Игнорирование жалоб граждан': 3 }").

Thus, this rule-based post-processing layer transforms raw probabilistic outputs into immediately actionable, human-readable competency profiles and personalised development recommendations, significantly enhancing practical utility and interpretability of the system.

The classification head consists of a single linear layer applied to the [CLS] token representation, followed by softmax activation. This produces probabilities for the two classes (`Is_Good` = 1 "good decision"; `Is_Good` = 0 "poor decision").

To evaluate the model, two metrics were employed – accuracy (0.986 8 in our case) and F1-score (0.987 0) calculated using the following formula:

$$F1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}},$$

where $\text{precision} = \frac{TP}{TP + FP}$; $\text{recall} = \frac{TP}{TP + FN}$; TP, FP and FN denote the number of true positives, false positives and false negatives, respectively.

The experiments utilised 5-fold stratified cross-validation with data shuffling and a fixed random state (`random_state=42`) to ensure reproducibility. The folds were stratified by the target class `Is_Good`, maintaining an equal proportion of labels 1 and 0 in each fold. This approach enables a reliable assessment of the model's generalisation ability and mitigates the risk of overfitting.

Results and discussion

On the independent test set (76 instances), the model achieved the following metrics: accuracy = 0.986 8, F1-score = 0.987 0. The confusion matrix (fig. 2) shows only one misclassification (one false positive), confirming the high accuracy of the model. The receiver operating characteristic (ROC) curve (fig. 3) has an area under the curve (AUC) of 0.998 6, significantly outperforming the random classifier line (0.5), which indicates excellent discriminative ability.

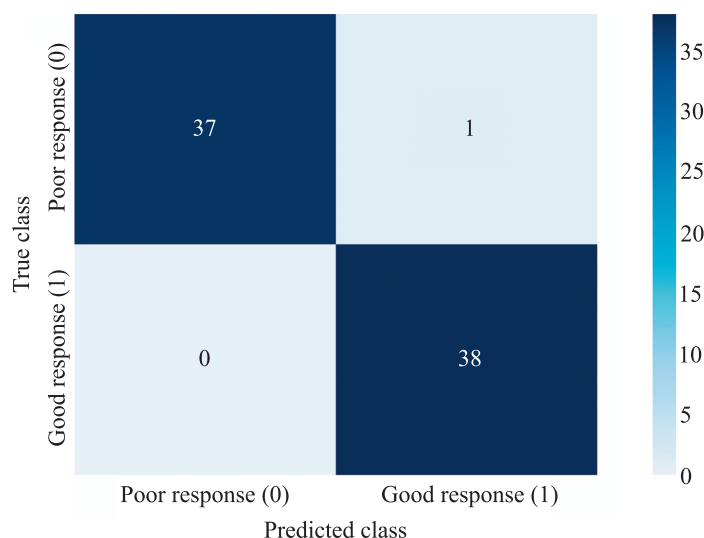


Fig. 2. Confusion matrix

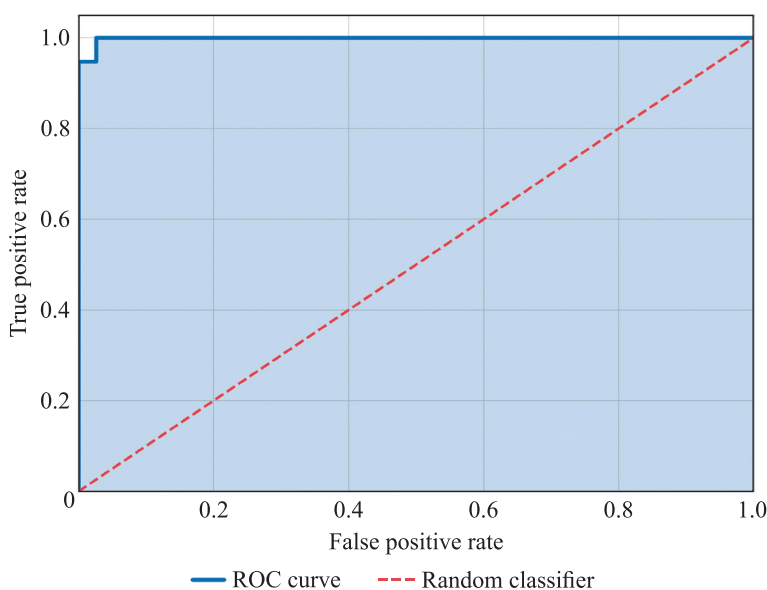


Fig. 3. ROC curve

The average metrics of the 5-fold stratified cross-validation are shown in the table 1. Detailed per-fold metrics are shown in fig. 4. The lowest performance is observed in the third fold (accuracy = 0.927 6, F1-score = 0.922 0), possibly due to the distribution of more challenging examples in that subset. Nevertheless, the overall stability across folds confirms the model’s strong generalisation capability.

Table 1

The average metrics
of the 5-fold stratified cross-validation

Metrics	Value
Accuracy	0.9711 ± 0.0230
F1-score	0.9700 ± 0.0252
Precision	0.9819 ± 0.0130
Recall	0.9605 ± 0.0546

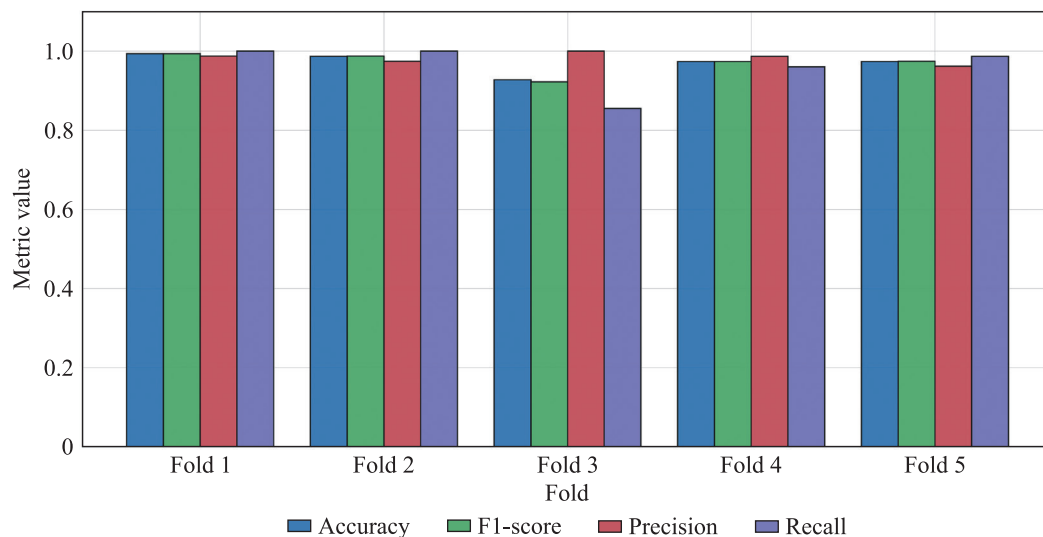


Fig. 4. Metrics across each fold

The histogram of predicted probabilities for the «Good response» class (fig. 5) demonstrates clear class separation: most good responses received probabilities >0.95 , while poor responses received probabilities <0.05 . This indicates high model confidence in predictions and a low risk of ambiguous classifications.

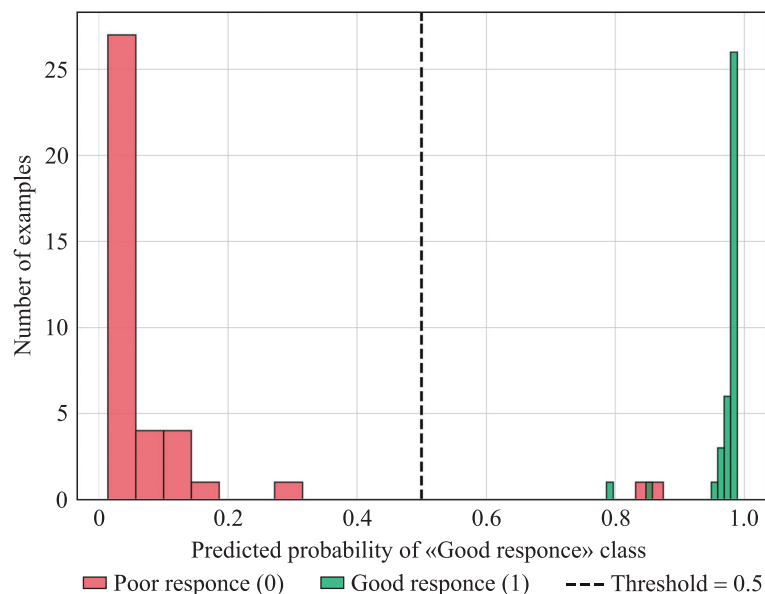


Fig. 5. Distribution of predicted probabilities

Example of model inference for Case_ID 71 (good response)

Input: "В районе участились случаи граффити на общественных зданиях. [SEP] Организую конкурс уличного искусства, выделяю легальные стены и запущу профилактическую работу с подростками."

Model output: Is_Good = 1 (probability 0.987)

Interpretation: "Good response. Competencies demonstrated: Профилактика и долгосрочный подход (5), Инновационность и оптимизация процессов (5)."

Table 2

Test set results

Metrics	Value
Accuracy	0.986 8
F1-score	0.987 0
Precision	0.974 4
Recall	1.000 0
ROC-AUC	0.998 6

The obtained metrics significantly exceed typical values for contextual text classification tasks in Russian and approach the upper limits achievable on synthetic balanced datasets. The high accuracy (above 0.97 even in cross-validation) validates the effectiveness of the proposed context-aware approach with concatenation of the situation and response. The deterministic post-processing of predictions adds practical value, transforming binary classification into a tool for generating personalised competency-based feedback. The slight drop in the third fold highlights a potential improvement area – possibly through data augmentation or model ensembling.

Comparison with existing solutions

The task could also be addressed using the full-sized BERT model. While it may yield slightly higher accuracy, it demands significantly greater computational resources. Due to its larger parameter count (110 mln versus 66 mln in DistilBERT), both training time and memory consumption are substantially higher.

Modified BERT variants, such as RoBERTa (employs improved pre-training objectives) or ALBERT (relies on parameter factorisation and sharing), can offer competitive or superior performance. However, these models are more complex to fine-tune and generally more resource-intensive than DistilBERT [25; 26].

Traditional machine learning approaches, including support vector machines, random forest and naive Bayes, remain viable for text classification but require extensive manual feature engineering (e. g., TF-IDF vectorisation). In practice, these methods typically underperform transformer-based models (especially on complex contextual tasks such as evaluating managerial responses) due to their limited ability to capture long-range dependencies and semantic nuances [27; 28].

The key advantages of the proposed approach are as follows:

1) DistilBERT is approximately 45 % faster and 40 % lighter than BERT while retaining over 96 % of its performance, making it far more suitable for real-world deployment;

2) the model not only performs binary classification (Is_Good = 0 or Is_Good = 1) but also extracts demonstrated competencies and critical red flags, thereby delivering actionable insights highly valuable to HR and training departments – an output rarely provided by standard classification models;

3) the achieved performance metrics (accuracy = 0.986 8, F1-score = 0.987 0, with cross-validation showing accuracy = 0.9711 ± 0.0230 and F1-score = 0.9700 ± 0.0252 across 5 folds) are among the highest reported for this type of task and surpass numerous results obtained with DistilBERT in the literature [29] (for instance, reference [22] reports a maximum accuracy of 99.41 % on certain benchmark datasets, which our solution approaches on a significantly more contextually complex managerial evaluation task);

4) concatenating the situational description and candidate response as a single input sequence (Situation + Response) enables the model to fully account for context, a factor of critical importance in assessing the quality of managerial decisions.

Among the limitations of the proposed solution are the requirement for a large volume of high-quality labelled training data and the maximum sequence length constraint of 512 tokens inherent to BERT-based architectures.

Conclusions

The present study demonstrates the high effectiveness of a fine-tuned DistilBERT model for the automatic evaluation of managerial decisions based on textual responses paired with situational descriptions. A key contribution is the creation of a specialised Russian-language dataset enriched with expert annotations of

demonstrated competencies and developmental areas (red flags). By incorporating a post-processing layer that maps model predictions to specific competency scores (e. g., 'Community engagement' : 5) and critical shortcomings (e. g., 'Inaction'), the proposed solution transcends conventional binary classification and significantly enhances interpretability and practical utility of the results.

The achieved performance metrics (accuracy = 0.986 8, F1-score = 0.987 0) with robust cross-validation results (mean accuracy = 0.971 1 $\pm$ 0.023 0) establishes a strong benchmark and opens promising avenues for deploying automated decision-support systems in management training, personnel assessment and HR analytics. Future research directions include:

- 1) augmenting the original dataset with real-world managerial cases to improve ecological validity;
- 2) benchmarking other language models (including multi-lingual and larger Russian-specific models) on the proposed dataset;
- 3) extending the competency framework and red flags taxonomy to cover a broader range of managerial and leadership skills.

References

1. Hinton G, Vinyals O, Dean J. Distilling the knowledge in a neural network. arXiv:1503.02531 [Preprint]. 2015 [cited 2024 December 19]: [9 p.]. Available from: <https://arxiv.org/abs/1503.02531>.
2. Prema V, Elavazhahan V. Sculpting DistilBERT: enhancing efficiency in resource-constrained scenarios. In: Dwivedi RK, Saxena AK, Sharma R, Bhardwaj S, Khattri V, editors. *Proceedings of the 2023 12th International conference on system modeling & advancement in research trends (SMART); 2023 December 22–23; Moradabad, India*. [S. l.]: Institute of Electrical and Electronics Engineers; 2023. p. 251–256. DOI: 10.1109/SMART59791.2023.10428568.
3. Mukherjee A, Umme Salma M. Resume ranking and shortlisting with DistilBERT and XLM. In: Institute of Electrical and Electronics Engineers. *Proceedings of the 2024 IEEE International conference for women in innovation, technology and entrepreneurship (ICWITE); 2024 February 16–17; Bangalore, India*. [S. l.]: Institute of Electrical and Electronics Engineers; 2024. p. 301–304. DOI: 10.1109/ICWITE59797.2024.10502523.
4. Jingga K, Hidayatullah H. The utilization of DistilBERT features space for topics suggestion in the context of learning content. In: Jusman Y, Mat Isa NA, Sukamta, Ejaz W, Ismail AH, Bhattacharjya A, et al., editors. *Proceedings of the 2023 International workshop on artificial intelligence and image processing (IWAIP); 2023 December 1–2; Yogyakarta, Indonesia*. [S. l.]: Institute of Electrical and Electronics Engineers; 2023. p. 243–247. DOI: 10.1109/IWAIP58158.2023.10462787.
5. Islam MT, Parvin F, Sazan SA, Amir TB. Comparative analysis of sentiment classification on IMDB 50k movie reviews: a study using CNN, LSTM, CNN-LSTM, and BERT models. In: Institute of Electrical and Electronics Engineers. *Proceedings of the 2024 IEEE International conference on power, electrical, electronics and industrial applications (PEELACON); 2024 September 12–13; Rajshahi, Bangladesh*. [S. l.]: Institute of Electrical and Electronics Engineers; 2024. p. 512–517. DOI: 10.1109/PEELACON63629.2024.10800035.
6. Otani N, Bhutani N, Hruschka E. Natural language processing for human resources: a survey. In: Chen W, Yang Y, Kachuee M, Fu X-Y, editors. *Proceedings of the 2025 Conference of the nations of the Americas chapter of the Association for Computational Linguistics: human language technologies; April 2025; Albuquerque, USA. Volume 3, Industry track*. Kerrville: Association for Computational Linguistics; 2025. p. 583–597. DOI: 10.18653/v1/2025.naacl-industry.47.
7. Kang Y, Cai Z, Tan C-W, Huang Q, Liu H. Natural language processing (NLP) in management research: a literature review. *Journal of Management Analytics*. 2020;7(2):139–172. DOI: 10.1080/23270012.2020.1756939.
8. Schmiedel T, Müller O, vom Brocke J. Topic modeling as a strategy of inquiry in organizational research: a tutorial with an application example on organizational culture. *Organizational Research Methods*. 2019;22(4):941–968. DOI: 10.1177/1094428118773858.
9. Jia J, Liang W, Liang Y. A review of hybrid and ensemble in deep learning for natural language processing. arXiv:2312.05589v2 [Preprint]. 2024 [cited 2024 December 19]: [23 p.]. Available from: <https://arxiv.org/abs/2312.05589v2>.
10. Davenport T, Guha A, Grewal D, Bressgott T. How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*. 2020;48(1):24–42. DOI: 10.1007/s11747-019-00696-0.
11. Golestani A, Masli M, Shami NS, Jones J, Menon A, Mondal J. Real-time prediction of employee engagement using social media and text mining. In: Wani MA, Kantardzic M, Sayed-Mouchaweh M, Gama J, Lughofer E, editors. *Proceedings of the 17th IEEE International conference on machine learning and applications (ICMLA); 2018 December 17–20; Orlando, USA*. [S. l.]: Institute of Electrical and Electronics Engineers; 2018. p. 1383–1387. DOI: 10.1109/ICMLA.2018.00225.
12. Jamuna KM. Leveraging natural language processing to analyze employee feedback for enhanced HR insights. *International Journal of Scientific Research in Engineering and Management*. 2024;8(11):1–6. DOI: 10.55041/IJSREM39115.
13. Leidner JL, Stevenson M. Challenges and opportunities of NLP for HR applications: a discussion paper. arXiv:2405.07766 [Preprint]. 2024 [cited 2024 December 19]: [10 p.]. Available from: <https://arxiv.org/abs/2405.07766>.
14. Zhang M, Jensen KN, Sonniks SD, Plank B. SkillSpan: hard and soft skill extraction from English job postings. In: *Proceedings of the 2022 Conference of the North American chapter of the Association for Computational Linguistics: human language technologies; 2022 July 10–15; Seattle, USA*. Stroudsburg: Association for Computational Linguistics; 2022. p. 4962–4984. DOI: 10.18653/v1/2022.naacl-main.366.
15. Uma VR, Velchamy I, Upadhyay D. Recruitment analytics: hiring in the era of artificial intelligence. In: Tyagi P, Chilamkurti N, Grima S, Sood K, Balusamy B, editors. *The adoption and effect of artificial intelligence on human resources management. Part A*. Bingley: Emerald Publishing Limited; 2023. p. 155–174 (Özen E, Grima S, editors. Emerald studies in finance, insurance, and risk management; volume 7).
16. França TJF, Mamede HS, Barroso JMP, dos Santos VMPD. Artificial intelligence applied to potential assessment and talent identification in an organisational context. *Heliyon*. 2023;9(4):e14694. DOI: 10.1016/j.heliyon.2023.e14694.

17. Devlin J, Chang M-W, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. In: Burstein J, Doran C, Solorio T, Kumar R, Loukina A, Morales M, et al., editors. *Proceedings of the 2019 Conference of the North American chapter of the Association for Computational Linguistics: human language technologies; 2019 June 2–7; Minneapolis, USA. Volume 1, Long and short papers*. Stroudsburg: Association for Computational Linguistics; 2019. p. 4171–4186. DOI: 10.18653/v1/N19-1423.
18. Gardazi NM, Daud A, Malik MK, Bukhari A, Alsahfi T, Alshemaimri B. BERT applications in natural language processing: a review. *Artificial Intelligence Review*. 2025;58(6):166. DOI: 10.1007/s10462-025-11162-5.
19. Davoodi L, Mezei J, Heikkilä M. Aspect-based sentiment classification of user reviews to understand customer satisfaction of e-commerce platforms. *Electronic Commerce Research*. 2026;26(2):1417–1459. DOI: 10.1007/s10660-025-09948-4.
20. Abdel-Salam S, Rafea A. Performance study on extractive text summarization using BERT models. *Information*. 2022;13(2):67. DOI: 10.3390/info13020067.
21. Lan Z, Chen M, Goodman S, Gimpel K, Sharma P, Soricut R. ALBERT: a lite BERT for self-supervised learning of language representations. arXiv:1909.11942v6 [Preprint]. 2020 [cited 2024 December 19]: [17 p.]. Available from: <https://arxiv.org/abs/1909.11942v6>.
22. Barbon RS, Akabane AT. Towards transfer learning techniques – BERT, DistilBERT, BERTimbau, and DistilBERTimbau for automatic text classification from different languages: a case study. *Sensors*. 2022;22(21):8184. DOI: 10.3390/s22218184.
23. Sanh V, Debut L, Chaumond J, Wolf T. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. arXiv: 1910.01108v4 [Preprint]. 2020 [cited 2024 December 19]: [5 p.]. Available from: <https://arxiv.org/abs/1910.01108>.
24. Andrenko KV, Kroshchanka AA, Golovko OV. The use of distilled large language models to determine the sentiment of a text. In: Golenkov VV, Azarov IS, Golovko VA, Gordey AN, Guliakina NA, Krasnoproshin VV, et al., editors. *Open semantic technologies for intelligent systems. Issue 9*. Minsk: Belarusian State University of Informatics and Radioelectronics; 2025. p. 229–234.
25. Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, et al. RoBERTa: a robustly optimized BERT pretraining approach. arXiv: 1907.11692 [Preprint]. 2019 [cited 2024 December 19]: [13 p.]. Available from: <https://arxiv.org/abs/1907.11692>.
26. Jiao X, Yin Y, Shang L, Jiang X, Chen X, Li L, et al. TinyBERT: distilling BERT for natural language understanding. In: Cohn T, He Y, Liu Y, editors. *Findings of the Association for Computational Linguistics: EMNLP-2020; 2020 November 16–20*. Stroudsburg: Association for Computational Linguistics; 2020. p. 4163–4174. DOI: 10.18653/v1/2020.findings-emnlp.372.
27. Golovko V, Kroshchanka A, Treadwell D. The nature of unsupervised learning in deep neural networks: a new understanding and novel approach. *Optical Memory and Neural Networks*. 2016;25(3):127–141. DOI: 10.3103/S1060992X16030073.
28. Golovko V, Kroshchanka A, Turchenko V, Jankowski S, Treadwell D. A new technique for restricted Boltzmann machine learning. In: Institute of Electrical and Electronics Engineers. *Proceedings of the 2015 IEEE 8th International conference on intelligent data acquisition and advanced computing systems: technology and applications (IDAACS); 2015 September 24–26; Warsaw, Poland. Volume 1*. [S. l.]: Institute of Electrical and Electronics Engineers; 2015. p. 182–186. DOI: 10.1109/IDAACS.2015.7340725.
29. Andrenko KV, Kroshchanka AA, Golovko VA, Krapivin YuB. The use of neural networks for decision-making analysis. In: Ablameyko SV, Kazachonak VU, Kurbackij AN, Krasnoproshin VV, editors. *Information systems and technologies. Proceedings of the 11th International scientific congress on computer science (CSIST-2025); 2025 October 29–31; Minsk, Belarus. Part 2*. Minsk: Belarusian State University; 2025. p. 118–125. Russian.

Received 24.12.2025 / revised 19.01.2026 / accepted 29.01.2026.

СИСТЕМА РАСПОЗНАВАНИЯ ПОХОДКИ В РЕАЛЬНОМ ВРЕМЕНИ НА ОСНОВЕ ЛЕГКОВЕСНОЙ ДВУХМОДАЛЬНОЙ ОЦЕНКИ ПОЛОЖЕНИЯ

ДИН АОДИ¹⁾, ЙЕ ШИПИНГ²⁾, А. М. НЕДЗЬВЕДЬ^{1), 3)}, В. С. АНОСОВ⁴⁾

¹⁾Белорусский государственный университет, пр. Независимости, 4, 220030, г. Минск, Беларусь

²⁾Чжэцзян-Шуренский университет, ул. Шурен, 8, 310015, г. Ханчжоу, Китай

³⁾Объединенный институт проблем информатики НАН Беларуси,
ул. Сурганова, 6, 220012, г. Минск, Беларусь

⁴⁾6-я городская клиническая больница, ул. Уральская, 5, 220037, г. Минск, Беларусь

Аннотация. Предлагается облегченный подход к совместному анализу походки в 2D- и 3D-пространствах. Реализуется двухмодальная система распознавания походки, обеспечивающая ее высокоточную оценку в реальном времени за счет архитектуры совместной 2D- и 3D-реконструкции. В рамках системы для детекции ключевых точек скелета в 2D-пространстве и извлечения биометрических метрик походки применяется библиотека MediaPipe. Для трехмерной оценки позы используется модифицированная облегченная архитектура MTSA-former, которая содержит 2,5 млн параметров. Данная модель реализует каскадную обработку, включающую отдельные временные и пространственные модули, а также матрицу смежности, которая построена на основе топологической структуры скелета, что позволяет эффективно моделировать физические ограничения между суставами. Для достижения высокой производительности при сохранении вычислительной легкости предлагается гибридная локально-глобальная стратегия пространственного моделирования, интегрирующая графовые сверточные сети. Такой подход обеспечивает точную и эффективную реконструкцию 3D-позы в условиях реального времени.

Ключевые слова: анализ походки; оценка 3D-позы; облегченная модель; двухмодальное распознавание.

Благодарность. Работа была выполнена при финансовой поддержке Китайского стипендиального совета и Министерства трудовых ресурсов и социального обеспечения Китая (гранты № H20250551 и H20240330).

Образец цитирования:

Дин Аоди, Йе Шипинг, Недзьведь АМ, Аносов ВС. Система распознавания походки в реальном времени на основе легкой двухмодальной оценки положения. *Журнал Белорусского государственного университета. Математика. Информатика.* 2026;1:130–139 (на англ.).
EDN: XDQZRA

For citation:

Din Aodi, Ye Shiping, Nedzved AM, Anosov VS. Real-time gait recognition system based on lightweight dual-modal pose estimation. *Journal of the Belarusian State University. Mathematics and Informatics.* 2026;1:130–139.
EDN: XDQZRA

Авторы:

Дин Аоди – аспирант кафедры информационных систем управления факультета прикладной математики и информатики. Научный руководитель – А. М. Недзьведь.

Йе Шипинг – проректор.

Александр Михайлович Недзьведь – доктор технических наук, доцент; заведующий кафедрой информационных систем управления факультета прикладной математики и информатики¹⁾, главный научный сотрудник отдела интеллектуальных информационных систем³⁾.

Виктор Сергеевич Аносов – главный врач.

Authors:

Din Aodi, postgraduate student at the department of information management systems, faculty of applied mathematics and computer science.

Ye Shiping, vice-rector.

Alexander M. Nedzved, doctor of science (engineering), docent; head of the department of information management systems, faculty of applied mathematics and computer science^a, and chief researcher at the department of intellectual information systems^c.

Victor S. Anosov, chief physician.

Victor S. Anosov, chief physician.
avictor8@gmail.com

REAL-TIME GAIT RECOGNITION SYSTEM BASED ON LIGHTWEIGHT DUAL-MODAL POSE ESTIMATION

DIN AODI^a, YE SHIPING^b, A. M. NEDZVED^{a,c}, V. S. ANOSOV^d

^aBelarusian State University, 4 Niezaliezhnasci Avenue, Minsk 220030, Belarus

^bZhejiang Shuren University, 8 Shuren Street, Hangzhou 310015, China

^cUnited Institute of Informatics Problems, National Academy of Sciences of Belarus,
6 Surganava Street, Minsk 220012, Belarus

^d6th City Clinical Hospital, 5 Uralskaja Street, Minsk 220037, Belarus

Corresponding author: A. M. Nedzved (nedzveda@tut.by)

Abstract. A lightweight approach for joint 2D and 3D gait analysis is proposed. A dual-modal gait recognition system is adopted, achieving high-precision real-time analysis through a 2D and 3D collaborative estimation architecture. The system employs MediaPipe for real-time 2D keypoint detection and gait metric extraction. An improved lightweight architecture MTSA-former with 2.5 mln parameters is used for 3D pose estimation. This model utilises cascaded processing incorporating temporal and spatial modules, along with an adjacency matrix based on skeletal topological structure to model physical constraints between joints. A local-global hybrid spatial modelling strategy fusing graph convolutional networks is implemented to achieve efficient 3D pose estimation, attaining high performance while maintaining lightweight characteristics. This approach ensures accurate and efficient 3D pose reconstruction in real time.

Keywords: gait analysis; 3D pose estimation; lightweight model; dual-modal recognition.

Acknowledgements. This work was carried out with the financial support of the China Scholarship Council and the Ministry of Human Resources and Social Security of China (grants No. H20250551 and H20240330).

Introduction

Gait analysis, as a non-invasive method for motor function assessment and clinical diagnosis, holds significant application value in early screening. For example, Parkinson's disease patients exhibit typical characteristics such as shortened stride length, reduced gait speed and shuffling gait, while stroke patients present hemiplegic gait and limited joint range of motion [1]. These abnormal gait patterns can often be captured through quantitative analysis in the early stages of disease [2].

In fig. 1 the lightweight dual-modal gait recognition system proposed in this study, which balances real-time performance and accuracy through a 2D and 3D collaborative analysis architecture, is illustrated. MediaPipe first extracts 2D keypoint sequences from video in real-time, which are then input into the multi-scale temporal self-attention transformer model (here and further – MTSA-former model). This model captures gait cycles through multi-scale temporal convolution, reinforces skeletal constraints via spatial graph convolutional network (GCN) and ensures motion continuity through velocity and acceleration loss functions, thereby lifting 2D coordinates into high-precision 3D poses. Finally, based on the reconstructed 3D joint trajectories (particularly the lower limb hip, knee and ankle joints), end-to-end real-time gait analysis is achieved. Multi-dimensional gait features, including joint angle temporal curves, gait symmetry indices and time ratios of stance phase and swing phase, are extracted [3]. The lightweight design and dual-modal collaborative mechanism of this study enhance gait parameter measurement accuracy while maintaining computational efficiency [4]. Providing a novel technical approach for early warning of neurodegenerative diseases and quantitative assessment of rehabilitation outcomes.

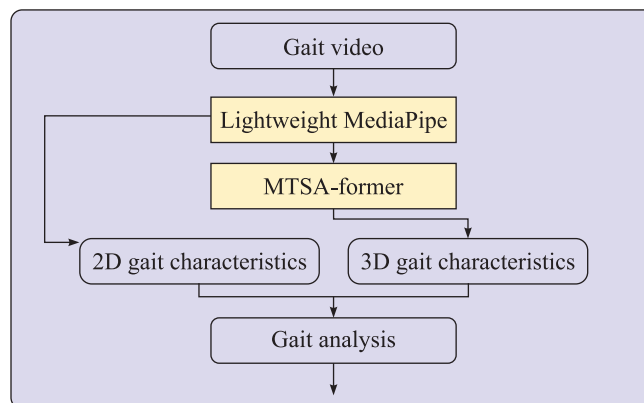


Fig. 1. Dual-modal gait recognition framework

Methods and principles

2D pose estimation method. MediaPipe utilises a lightweight detector to rapidly localise human body regions in images. During the pose keypoint localisation stage, the region of interest is cropped and normalised, then input into the pose estimation network to extract 2D coordinates of 22 human body keypoints. MediaPipe Lite model has approximately 3 mln parameters, achieving inference speeds of 30–60 frames per second on mobile central processing units and over 100 frames per second on graphics processing units [5]. With a model size of approximately 12 MB, it is suitable for deployment on resource-constrained edge devices. MediaPipe Lite significantly reduces computational cost and storage requirements while maintaining acceptable accuracy, providing an efficient 2D pose estimation foundation for the real-time dual-modal gait analysis system in this study [6].

From fig. 2, it can be seen that 11 skeletal nodes were removed from MediaPipe's original 33 keypoints, reducing the data dimensionality by approximately 33 %. The 22 core keypoints serve as the foundational skeletal model for gait analysis, encompassing the major movement joints of the human body, including key positions of the head, trunk, upper limbs and lower limbs. The most critical joints for gait analysis (lower limb joints (hip, knee and ankle joints) and trunk joints (shoulder and hip joints)) enable the model to more effectively capture key gait parameters such as periodic movement patterns during walking, joint angle variations and body center of gravity shifts.

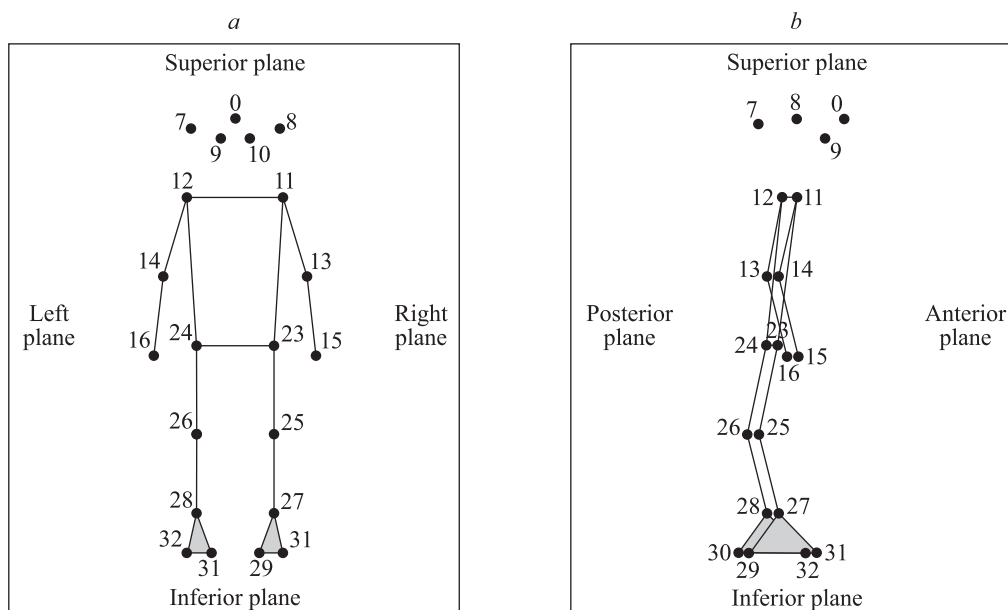


Fig. 2. MediaPipe key skeletal structure:
a – frontal view; b – sagittal view (left lateral)

3D pose estimation method. Traditional 3D human pose estimation models typically require large-scale annotated data, which poses challenges in clinical gait analysis scenarios due to high data collection costs for patients with gait disorders and limited sample sizes. This paper proposes MTSA-former model – a lightweight, highly efficient and rapidly convergent gait analysis-specific model trained on small samples from Human3.6M dataset, achieving fast training and accurate prediction under few-shot settings (fig. 3) [7]. The model employs depth-wise separable convolution, linear complexity attention mechanisms and parameter sharing strategies to significantly reduce model size while maintaining performance. The core of the transformer uses two main parts: attention and feed-forward network (FFN). The standard GELU function is used as the activation function.

The model first performs linear projection and positional encoding on the input sequence to obtain spatio-temporal feature embeddings. Subsequently, within multiple encoding blocks, temporal modelling (utilising multi-scale convolution to capture short-term and long-term motion variations) and spatial modelling (learning structural relationships between joints through graph convolution and attention mechanisms) are conducted simultaneously. Finally, multi-scale pyramid pooling fuses information from different temporal ranges, and the regression head outputs the 3D human pose. This design effectively enhances the model's spatiotemporal perception capabilities and prediction accuracy, demonstrating excellent performance in motion continuity and smoothness.

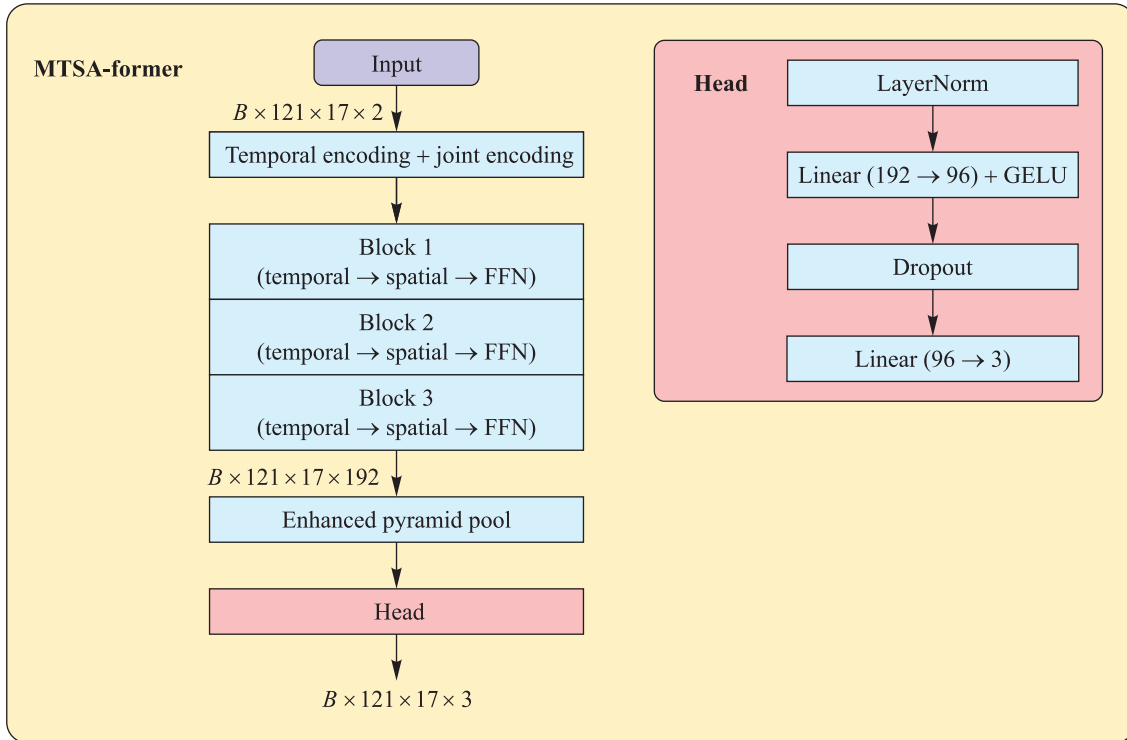


Fig. 3. 3D pose estimation framework using the MTSA-former model

Human skeletal spatial graph convolution. The human skeletal structure can be abstracted as a graph, where joint points are the nodes of the graph and bones are the edges connecting the nodes. GCN can explicitly model the physical connections and kinematic constraints between joint points by propagating and aggregating information across the graph [8]. Each GCN layer updates the feature representation of each node, where the new feature is a weighted sum of its own features and those of neighbouring nodes. By stacking GCN layers, the model can learn high-order inter-joint dependencies that conform to human structural priors, thereby effectively enhancing the physical realism and structural integrity of the generated poses. The composite loss function formula is as follows:

$$H^{(l+1)} = \sigma \left(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^{(l)} W^{(l)} \right).$$

Describes the feature propagation and transformation processes in a GCN from layer l to layer $l + 1$. Here $H^{(l)}$ denotes the node feature matrix at layer l ; $\tilde{A}$ represents the adjacency matrix; $\tilde{D}$ is the corresponding degree matrix; $W^{(l)}$ represents the trainable weight matrix at layer l .

The model is guided to learn a more comprehensive mapping relationship. It not only learns how to accurately localise the keypoints but also acquires prior knowledge of human body structure and the dynamic patterns of motion, ultimately generating highly realistic 3D human poses across spatial, structural and temporal dimensions. In formula

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{MPJPE}} + \alpha \mathcal{L}_{\text{bone}} + \gamma \mathcal{L}_{\text{velocity}} + \delta \mathcal{L}_{\text{acceleration}}$$

$\mathcal{L}_{\text{MPJPE}}$ measures the average Euclidean distance between the predicted 3D joint positions and the ground-truth joint positions; $\mathcal{L}_{\text{bone}}$ constrains the predicted pose to maintain bone lengths consistent with those of the real human body, preventing unrealistic stretching or compression; $\mathcal{L}_{\text{velocity}}$ is used to smooth the motion of joints between consecutive frames by penalising abrupt temporal changes, thereby ensuring smooth and continuous motion.

The 3D human pose estimation results of the MTSA-former model. As shown in fig. 4, the image on the left depicts the input 2D human keypoint detection, while the image on the right illustrates the reconstructed 3D human skeleton predicted by the model. This method can maintain high 3D reconstruction accuracy in gait detection scenarios, generating pose estimation results with strong structural consistency and smooth motion. It demonstrates the effectiveness of the improved model in multi-scale spatiotemporal feature fusion.

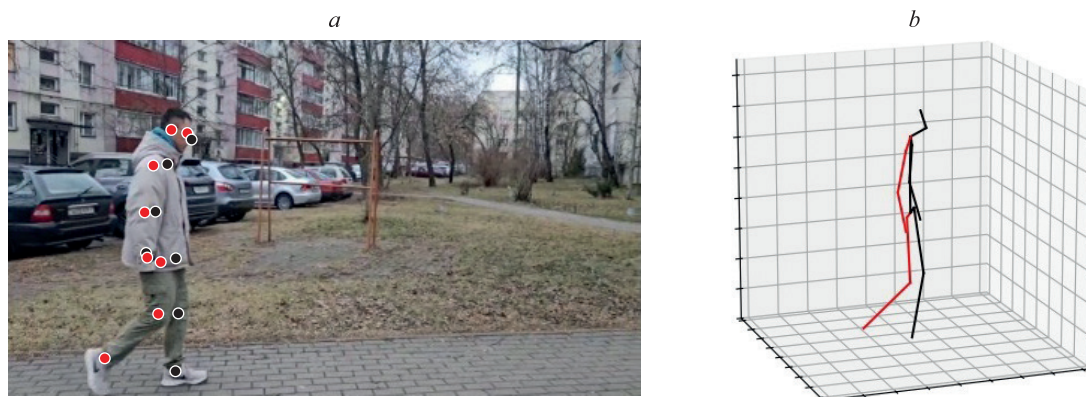


Fig. 4. 3D pose estimation using the MTSA-former model:
a – input 2D human keypoint detection;
b – reconstructed 3D human skeleton predicted by the model

Gait analysis techniques. Both 2D and 3D gait analysis are essential tools for evaluating gait and identifying deviations or abnormalities in movement. Gait analysis plays a vital role in the diagnosis and treatment of various musculoskeletal and neurological disorders. Using the logical framework motion features – normal range and abnormality grading – clinical significance, the angular variation parameters commonly observed during gait are systematically categorised and interpreted [9]. As shown in the table, the maximum ankle dorsiflexion angle, peak knee flexion during the swing phase and pelvic obliquity at midstance represent the core lowerlimb motion patterns in the 2D plane, whereas the knee progression angle and maximum trunk lateral shift reflect integrated manifestations of coordination and balance control in 3D space [10].

Gait analysis features

Feature	Expression dimension	Involved plane	Normal range	Mild abnormality	Moderate abnormality	Severe abnormality	Clinical significance
Maximum ankle dorsiflexion in stance	2D single plane	Sagittal plane	Dorsiflexion by 10–15°	Dorsiflexion by 5–9° or 16–20°	Dorsiflexion by 0–4° or 21–30°	Plantarflexion position or dorsiflexion by >30°	Assess ankle range of motion and gastrocnemius tightness
Peak knee flexion in swing	2D single plane	Sagittal plane	Flexion by 55–65°	Flexion by 45–54° or 66–75°	Flexion by 35–44° or 76–85°	Flexion by <35° or >85°	Assess foot clearance ability and quadriceps strength
Pelvic obliquity at mid-stance	2D single plane	Coronal plane	Drop of stance side by 4–6°	Drop by 2–3° or 7–9°	Drop by 0–1° or 10–12°	Elevation or drop by >12°	Assess hip abductor strength and trunk stability
Knee progression angle at mid-stance	3D composite	Sagittal and transverse planes	Knee forward direction by $\pm 5^\circ$	External or internal rotation by 6–10°	External or internal rotation by 11–20°	Rotational deviation by >20°	Assess lower limb rotational alignment and knee joint stress
Maximum trunk lateral shift	3D composite	Coronal and transverse planes	Lateral shift toward stance side by 3–5 cm	2–3 or 6–7 cm	1–2 or 8–10 cm	<1 or >10 cm	Assess balance control and energy efficiency

In gait analysis these indicators use joint angular variations as the core quantitative diagnostic basis. By precisely measuring the angular range of specific joints during the gait cycle and comparing it with standard normal intervals, clinicians can quantitatively assess the severity of issues such as muscle weakness, joint contracture or neuromotor impairment. For example, an ankle dorsiflexion angle of less than 5° during the stance phase often indicates tightness of the gastrocnemius or soleus muscles, while a knee progression angle de-

viating more than $\pm 20^\circ$ suggests potential lowerlimb rotational malalignment or deviation of the knee joint's mechanical axis.

In summary, this table establishes an angular-based quantitative classification system for gait abnormalities and provides a unified, repeatable and comparable set of parametric references for gait assessment in both clinical practice and research. It holds significant value for the diagnosis and therapeutic evaluation of gait using 2D and 3D analysis frameworks.

Experimental results

The lightweight, small-sample MTSA-former model was trained for 2D-to-3D human pose estimation. In fig. 5 the variations and convergence trends of key loss metrics during the training of the improved MTSA-former model over 23 epochs are illustrated. Both training and validation of mean per joint position error (MPJPE) exhibit consistent convergence, with validation error decreasing from 90.42 mm to a minimum of 41.12 mm at 18th epoch, representing a 54.5 % improvement (see fig. 5, *a*). The total loss curve demonstrates stable optimisation with occasional fluctuations at 10th and 19th epochs, likely due to learning rate scheduling (see fig. 5, *b*). The epoch-to-epoch validation changes reveal that major improvements occur in early training phases (see fig. 5, *c*), while fig. 5, *d*, shows that the best validation performance plateaus after 18th epoch, suggesting model convergence at this point. Overall, plots demonstrate a stable and well-converged training process for the 3D pose estimation model.

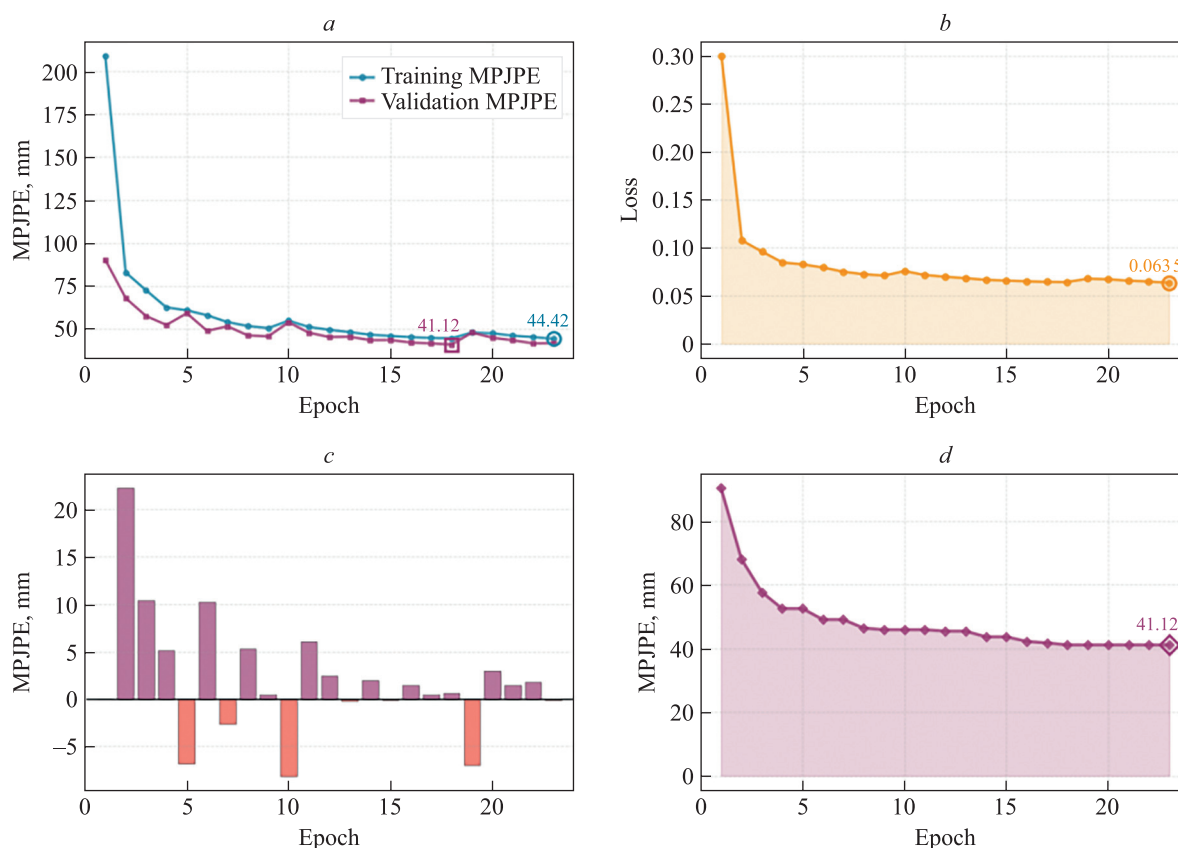


Fig. 5. Variations and convergence trends of key loss metrics during the training of the MTSA-former model:
a – training and validation MPJPE; *b* – total loss;
c – epoch-to-epoch validation change; *d* – best validation MPJPE progress

In fig. 6 the small-sample scenario on the Human3.6 dataset, reflecting the comparative relationship between model performance (measured by MPJPE) and model parameter size among various pose estimation models, based on a training set of 40 thsd samples and a test set of 8 thsd samples using the MTSA-former model framework is showed. The MTSA-former model achieves the lowest MPJPE of 41.12 mm with a relatively small number of parameters, demonstrating performance significantly superior to existing LSTM-based models such as PoseFormer and MHFormer [11; 12]. This indicates that the proposed method effectively reduces model complexity while maintaining high-precision pose estimation, showcasing a more optimised structural design and parameter utilisation efficiency.

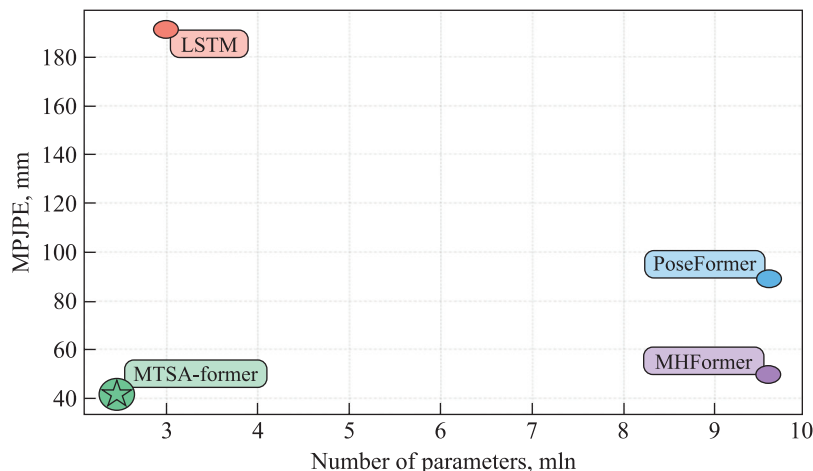


Fig. 6. Comparison of model parameters and accuracy.
The size of the circle corresponds to the performance score

The results of the comparative analysis of models LSTM, PoseFormer and MyModel, illustrating their MPJPE performance across 17 different joints, are presented in fig. 7. The comparison is used to evaluate variations and overall trends in spatial localisation accuracy for different parts of the human body. The proposed MyModel achieves the lowest MPJPE for all joints, significantly outperforming the baseline models LSTM and PoseFormer. This performance demonstrates that the model possesses clear advantages in local feature aggregation and global spatial structure preservation, enabling more accurate capture of the hierarchical relationships within human poses.

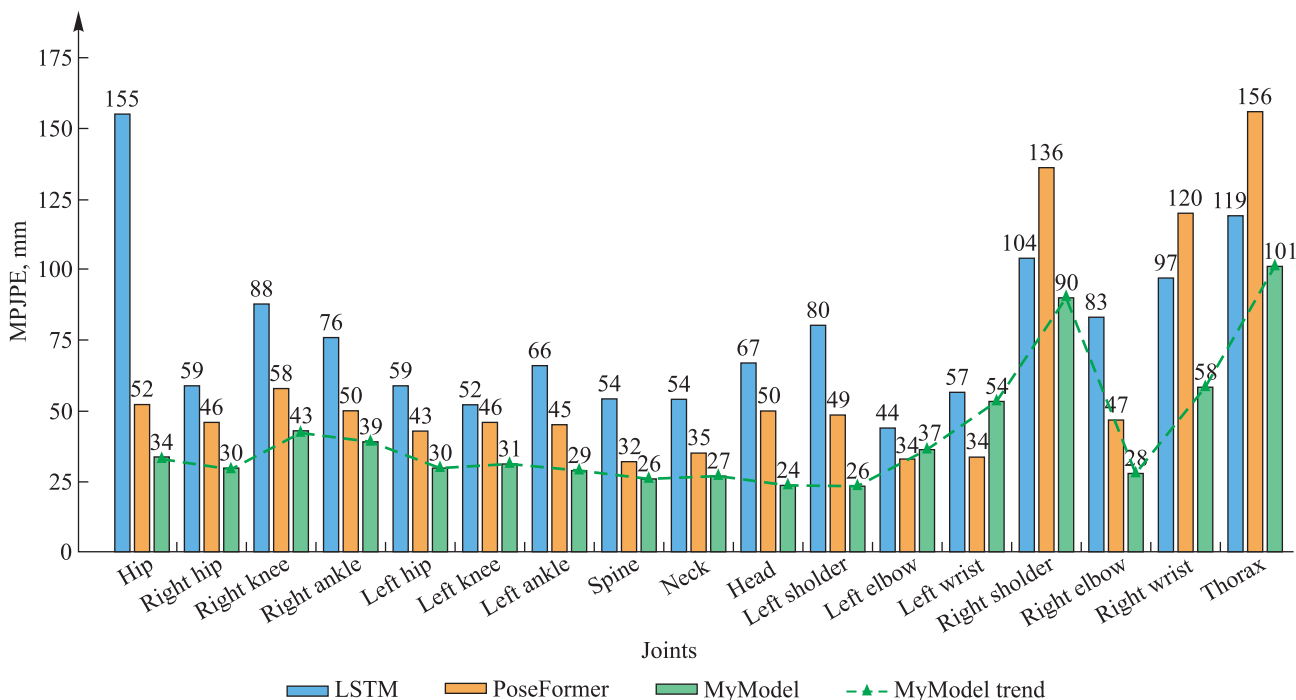


Fig. 7. Comparison of skeletal joints across different models

In gait analysis differences in key biomechanical parameters between healthy individuals and patients with Parkinson's disease are further examined (fig. 8). The core indicators, including foot-off events, ankle joint angles, trunk lateral shift and trunk forward inclination, are considered. The time-series analysis of foot-off height shows that the normal group maintains a stable average of 68.1 mm, whereas the Parkinson's disease group exhibits a markedly lower mean of 41.7 mm ($\Delta = 26.34$ mm), indicating reduced foot clearance and increased tripping risk (see fig. 8, a). The ankle joint angle distribution reveals a narrow, concentrated pattern in the normal group, while the Parkinson's disease group displays a broader, downward-shifted distribution (mean 151.4°), reflecting diminished mobility and increased rigidity (see fig. 8, b). Trunk lateral shift is minimal and smooth in

the normal group but increases significantly in the Parkinson's disease group (amplitude 31.8 mm), accompanied by high-frequency oscillations (evidence of impaired trunk stability and hip abductor weakness) (see fig. 8, *c*). The probability density of trunk sagittal inclination indicates a forward shift in the Parkinson's disease group (from -89.2° to -81.1°), with minimal overlap between groups, illustrating the characteristic stooped posture in Parkinson's gait (see fig. 8, *d*). The integrated four-dimensional radar plot, encompassing ankle angle, foot-off height, trunk lateral shift and trunk forward tilt, shows that the normal group forms a balanced, near-boundary diamond, whereas the Parkinson's disease group exhibits a contracted, asymmetric shape, visually highlighting multi-system gait impairment (see fig. 8, *e*).

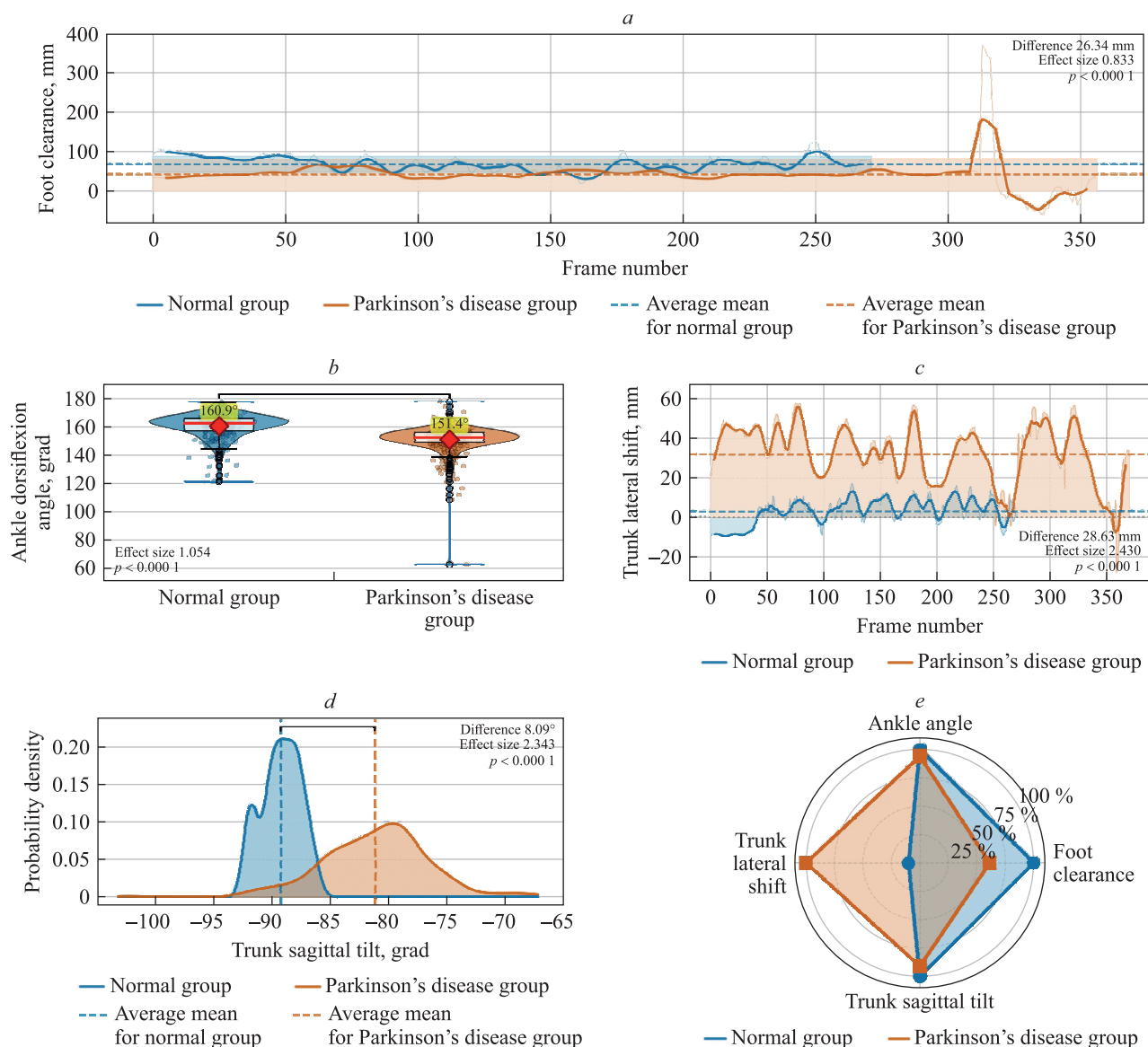


Fig. 8. Evaluation of gait in normal group and Parkinson's disease group:
a – foot clearance during swing phase; b – ankle angle;
c – trunk lateral displacement; d – trunk forward lean;
e – integrated four-dimensional radar plot

The visualisation results of a gait 2D motion analysis system implemented using the lightweight MediaPipe model, showcasing the capability of deep learning-based computer vision methods to perform real-time analysis of human dynamic behaviour under low computational cost conditions, is illustrated in fig. 9. The chart (see fig. 9) presents the dynamic waveform curves generated by the gait event analysis module, corresponding to the temporal motion variations of different body parts (such as the left heel and right toe). This system not only visualises the gait cycle but also enables the extraction of key gait features (including cadence, stride length and stance-to-swing ratio), providing valuable data support for subsequent applications in motor rehabilitation, behavioural monitoring and human biometric identification.

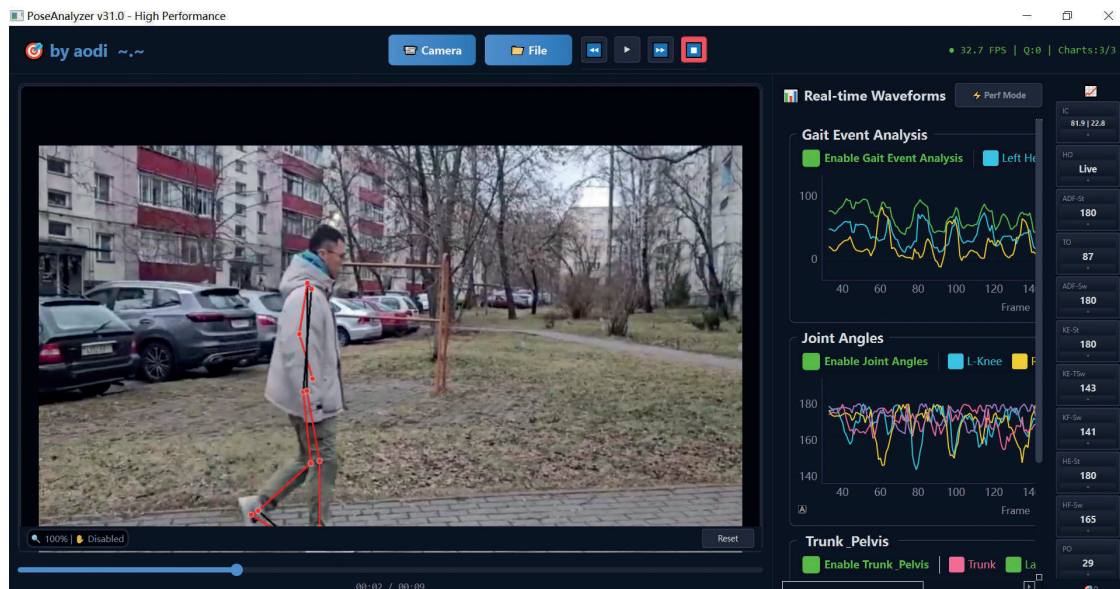


Fig. 9. Gait analysis system

A systematic academic analysis comparing healthy gait and Parkinsonian gait, incorporating the medical gait feature assessment table for quantitative inference, is provided in fig. 10. The analysis covers angular variations, motion plane characteristics and clinical significance, aiming to reveal differences in functional control and neuromuscular coordination between the two gait patterns. The normal's group left and right knee angle waveforms are showed mirror symmetry, stable periodicity (approximately one complete cycle per second) and an amplitude of $55\text{--}60^\circ$. The 3D kinematic characteristics fall within the normal physiological range, demonstrating symmetry, strong periodicity and good dynamic balance. In contrast, the Parkinson's disease group ankle, knee, hip and trunk exhibit are restricted ranges and markedly reduced amplitudes (around $30\text{--}35^\circ$). The gait cycle is irregular, and waveform peaks display noticeable disturbances. According to clinical assessment scales, this corresponds to a moderate-to-severe gait abnormality.

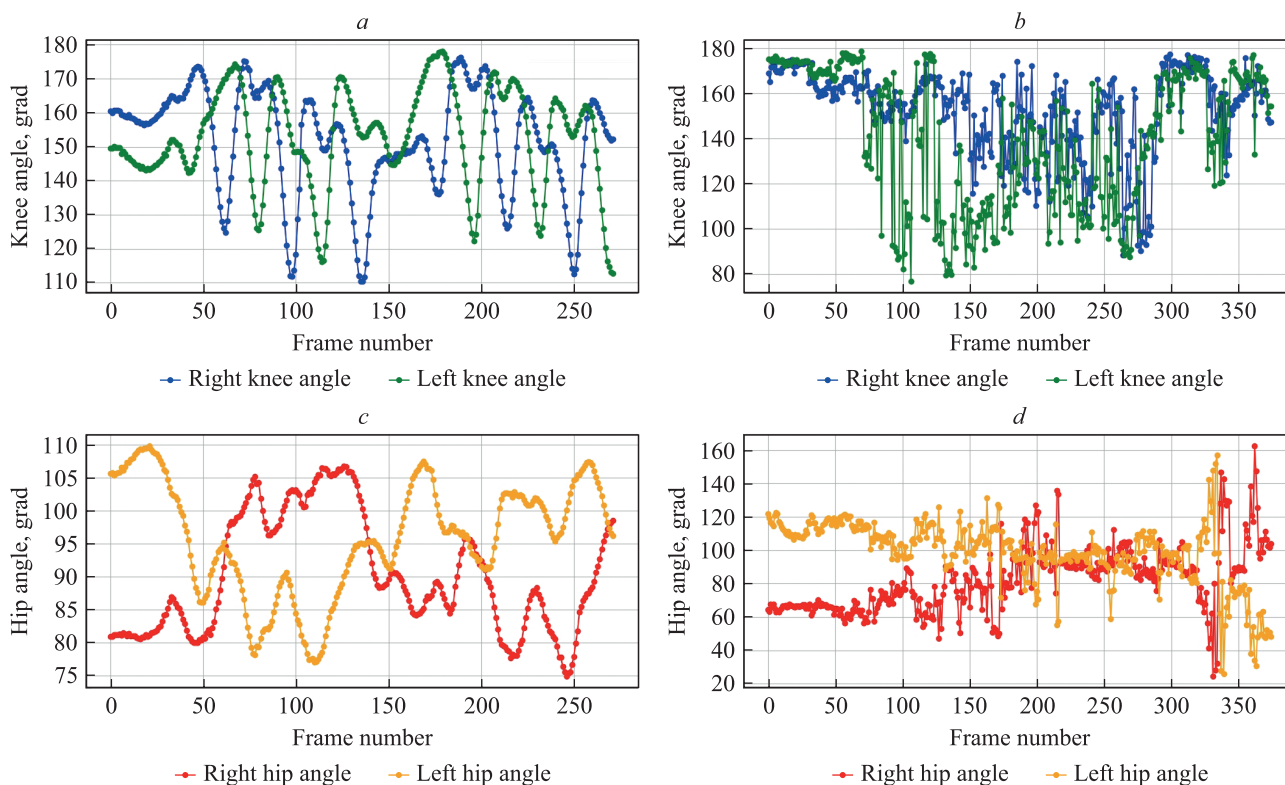


Fig. 10. 3D gait analysis, specifically knee joint angle changes (a, b), hip joint angle changes (c, d) of normal group (a, c) and Parkinson's disease group (b, d) (beginning)

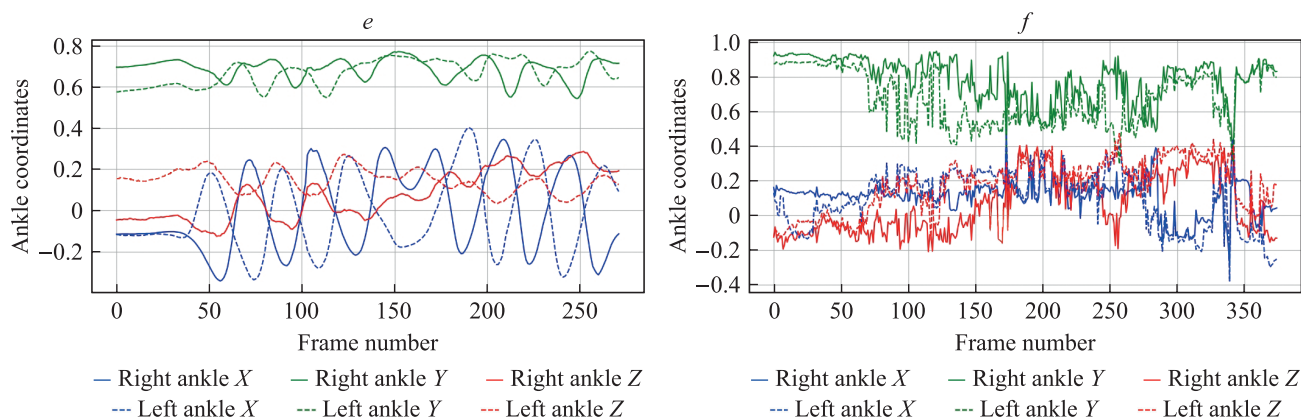


Fig. 10. 3D gait analysis, specifically ankle 3D coordinate changes, of normal group (e) and Parkinson's disease group (f) (ending)

Conclusions

This study presents a lightweight 2D and 3D collaborative gait analysis framework enabling high-precision, real-time modelling and quantitative clinical evaluation of human gait. The method integrates MediaPipe-based 2D keypoint detection with an enhanced MTSA-former 3D pose estimation network. Despite containing only 2.5 mln parameters, the model achieves superior 3D reconstruction accuracy (MPJPE = 41.12 mm) with real-time efficiency.

Comparative analyses of gait in normal group and Parkinson's disease group demonstrate the system's sensitivity in detecting abnormal variations in key biomechanical parameters, including ankle dorsiflexion, peak knee flexion, pelvic obliquity and trunk lateral shift. Clinical metrics further confirm its efficacy in distinguishing different motor impairment levels based on 2D kinematic and 3D spatiotemporal features. In gait of Parkinson's disease group the model identifies typical deficits such as limited ankle motion, inadequate knee flexion, reduced pelvic stability and impaired balance control.

References

1. Rastegari E, Azizian S, Ali H. Machine learning and similarity network approaches to support automatic classification of Parkinson's diseases using accelerometer-based gait analysis. In: Bui TX, editor. *Proceedings of the 52nd Hawaii international conference on system sciences*; 2019 January 8–11; Grand Wailea, USA. Honolulu: ScholarSpace; 2019. p. 4231–4242. DOI: 10.24251/HICSS.2019.511.
2. Tupa O, Procházka A, Vyšata O, Schätz M, Mařík J, Vališ M, et al. Motion tracking and gait feature estimation for recognizing Parkinson's disease using deep learning. *BioMedical Engineering OnLine*. 2015;14(1):97. DOI: 10.1186/s12938-015-0092-7.
3. Lugaresi C, Tang J, Nash H, McClanahan C, Uboweja E, Hays M, et al. MediaPipe: a framework for building perception pipelines. arXiv:1906.08172 [Preprint]. 2019 [cited 2025 August 24]: [9 p.]. Available from: <https://arxiv.org/pdf/1906.08172>.
4. Zeng W, Ouyang W, Luo P, Liu W, Wang X. 3D human mesh regression with dense correspondence. In: Institute of Electrical and Electronics Engineers. *Proceedings of the 2020 IEEE/CVF conference on computer vision and pattern recognition (CVPR)*; 2020 June 13–19; Seattle, USA. [S. l.]: Institute of Electrical and Electronics Engineers; 2020. p. 7052–7061. DOI: 10.1109/CVPR42600.2020.00708.
5. Bazarevsky V, Grishchenko I, Raveendran K, Zhu T, Zhang F, Grundmann M. BlazePose: on-device real-time body pose tracking. arXiv:2006.10204 [Preprint]. 2020 [cited 2025 May 22]: [4 p.]. Available from: <https://arxiv.org/pdf/2006.10204>.
6. Grishchenko I, Ablavatski A, Kartynnik Y, Raveendran K, Grundmann M. Attention mesh: high-fidelity face mesh prediction in real-time. arXiv:2006.10962 [Preprint]. 2021 [cited 2025 May 1]: [4 p.]. Available from: <https://arxiv.org/pdf/2006.10962>.
7. Ionescu C, Papava D, Olaru V, Sminchisescu C. Human3.6M: large scale datasets and predictive methods for 3D human sensing in natural environments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2014;36(7):1325–1339. DOI: 10.1109/TPAMI.2013.248.
8. Wandt B, Rosenhahn B. RepNet: weakly supervised training of an adversarial reprojection network for 3D human pose estimation. In: Institute of Electrical and Electronics Engineers. *Proceedings of the 2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR)*; 2019 June 15–20; Long Beach, USA. [S. l.]: Institute of Electrical and Electronics Engineers; 2019. p. 7774–7783. DOI: 10.1109/CVPR.2019.00797.
9. Baker R, McGinley JL, Schwartz MH, Beynon S, Rozumalski A, Graham HK, et al. The gait profile score and movement analysis profile. *Gait & Posture*. 2009;30(3):265–269. DOI: 10.1016/j.gaitpost.2009.05.020.
10. Kainz H, Carty ChP, Modenese L, Boyd RN, Lloyd DG. Estimation of the hip joint centre in human motion analysis: a systematic review. *Clinical Biomechanics*. 2020;30(4):319–329. DOI: 10.1016/j.clinbiomech.2015.02.005.
11. Zheng C, Zhu S, Mendieta M, Yang T, Chen C, Ding Z. 3D human pose estimation with spatial and temporal transformers. In: Institute of Electrical and Electronics Engineers. *Proceedings of the 2021 IEEE/CVF International conference on computer vision (ICCV)*; 2021 October 10–17; Montreal, Canada. [S. l.]: Institute of Electrical and Electronics Engineers; 2021. p. 11636–11645. DOI: 10.1109/ICCV48922.2021.01145.
12. Li W, Liu H, Tang H, Wang P, Van Gool L. MHFormer: multi-hypothesis transformer for 3D human pose estimation. In: Institute of Electrical and Electronics Engineers. *Proceedings of the 2022 IEEE/CVF conference on computer vision and pattern recognition (CVPR)*; 2022 June 18–24; New Orleans, USA. [S. l.]: Institute of Electrical and Electronics Engineers; 2022. p. 13147–13156.

Received 24.12.2025 / revised 19.01.2026 / accepted 29.01.2026.

МУЛЬТИМОДЕЛЬНЫЙ ПОДХОД С ВЕКТОРНЫМ ПОИСКОМ В ОПЕРАТИВНОМ УПРАВЛЕНИИ КРИТИЧНЫМ ИТ-СЕРВИСОМ

В. В. КРАСНОПРОШИН¹⁾, А. А. СТАРОВОЙТОВ¹⁾

¹⁾Белорусский государственный университет, пр. Независимости, 4, 220030, г. Минск, Беларусь

Аннотация. Исследуются актуальные проблемы, связанные с задачей оперативного управления критичными ИТ-сервисами. Для повышения качества управления предлагается новый метод, расширяющий подход динамической локальной аппроксимации. Он реализует связывание векторных представлений паттернов нагрузки с обучаемыми нейросетевыми моделями, что позволяет оперативно строить предиктор для нового паттерна на основе векторного поиска семантически близких моделей в библиотеке. Излагаются оригинальные принципы проектирования системы управления, которая для адаптации к новым, связанным с неопределенностью паттернам нагрузки использует аппроксимацию данных, получаемых в режиме реального времени. Приводятся результаты экспериментов, подтверждающие эффективность предложенного подхода и показывающие, что при обработке нерегулярных и сложных нагрузок гибридные проактивные алгоритмы по ряду показателей являются более эффективными, чем реактивные алгоритмы.

Ключевые слова: принятие решений; проактивное управление; неопределенность внешней нагрузки; нейронная сеть; мультимодельная система; библиотека моделей; векторный поиск.

MULTI-MODEL APPROACH WITH VECTOR SEARCH IN OPERATIONAL MANAGEMENT CRITICAL IT SERVICE

V. V. KRASNOPROSHIN^a, A. A. STAROVOITOV^a

^aBelarusian State University, 4 Niezaliezhnasci Avenue, Minsk 220030, Belarus

Corresponding author: V. V. Krasnoproshin (krasnoproshin@bsu.by)

Abstract. In the work are investigated current problems related to the task of operational management of critical IT services. To improve the quality of management, a new method is proposed, extending the dynamic local approximation approach. It implements linking vector representations of load patterns with trainable neural network models, which allows promptly constructing a predictor for a new pattern based on vector search for semantically close models in the library. Original principles of designing a control system are outlined, which for adaptation to new, uncertainty-related load patterns uses approximation of data obtained in real-time mode. The results of experiments confirmed the effectiveness

Образец цитирования:

Краснопрошин ВВ, Старовойтов АА. Мультимодельный подход с векторным поиском в оперативном управлении критичным ИТ-сервисом. *Журнал Белорусского государственного университета. Математика. Информатика.* 2026; 1:140–155 (на англ.).
EDN: OYKWAO

For citation:

Krasnoproshin VV, Starovoitov AA. Multi-model approach with vector search in operational management critical IT service. *Journal of the Belarusian State University. Mathematics and Informatics.* 2026;1:140–155.
EDN: OYKWAO

Авторы:

Виктор Владимирович Краснопрошин – доктор технических наук, профессор; профессор кафедры информационных систем управления факультета прикладной математики и информатики.
Александр Александрович Старовойтов – аспирант кафедры информационных систем управления факультета прикладной математики и информатики. Научный руководитель – В. В. Краснопрошин.

Authors:

Viktor V. Krasnoproshin, doctor of science (engineering), full professor; professor at the department of information management systems, faculty of applied mathematics and computer science.
krasnoproshin@bsu.by
<https://orcid.org/0000-0002-9463-4869>
Aleksandr A. Starovoitov, postgraduate student at the department of information management systems, faculty of applied mathematics and computer science.
starovoytova@bsu.by
<https://orcid.org/0009-0001-6531-4225>

of the proposed approach and showed that when processing irregular and complex loads, hybrid proactive algorithms in a number of indicators are more effective than reactive ones.

Keywords: decision making; proactive control; external load uncertainty; neural network; multi-model system; model library; vector search.

Introduction

Maintaining critical IT services in operational condition with a guaranteed level of service (*service level objective*, SLO) is one of the key tasks of the operation of modern digital infrastructure. The dynamic and often unpredictable nature of external load requires the use of complex automatic scaling (auto-scaling) mechanisms, capable not only of reacting to sharp load changes but also of predicting them [1].

Existing industrial solutions (Kubernetes HPA, AWS Auto Scaling, Azure Autoscale) are often based on threshold rules, not always effective under complex, non-stationary load profiles. Research approaches, using machine learning and time series forecasting, show high accuracy, but may require long training and are not sufficiently prompt in adapting to new, previously unseen load patterns. As an example illustrating the complexity of the load profile in fig. 1 are presented data on user requests to the web server of the NASA Kennedy Space Centre in Florida on 6 July 1995¹.

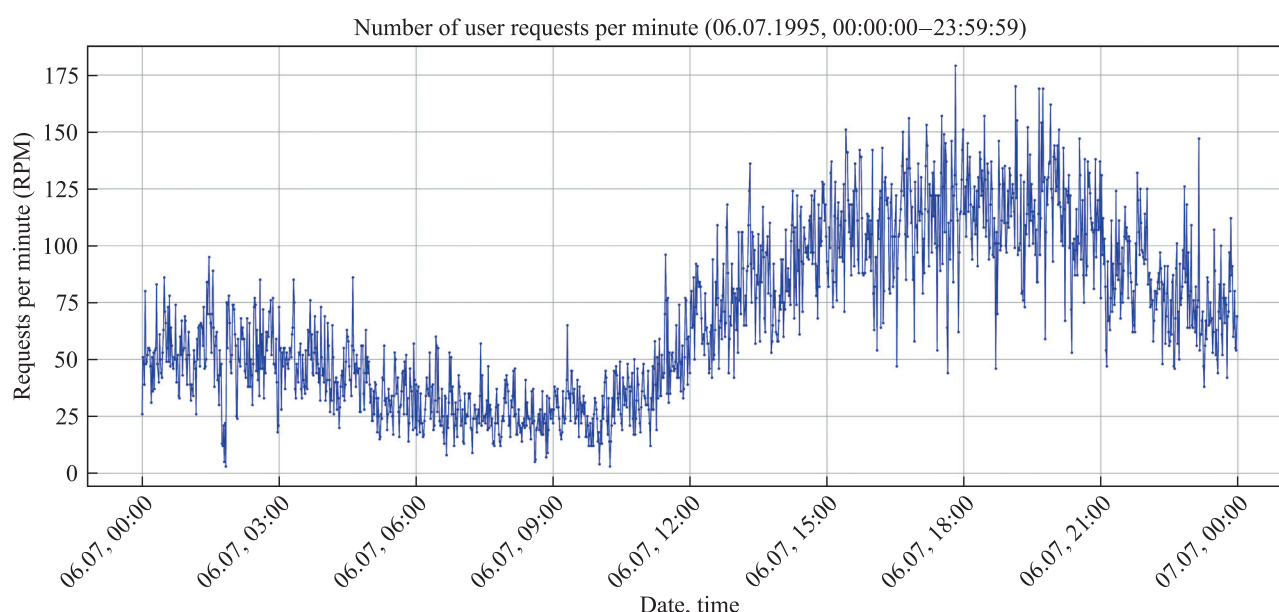


Fig. 1. Example of external load

In previous work [2], the authors proposed a multi-model method DLANNLIB (*dynamic local approximation by neural network models with library*). For each system state (determined by the volume of allocated resources) in real time is built and saved into a library a local neural network model, trained on data of the current load pattern. Such a mechanism allows accumulating «experience» and using it for predicting the transition of the model system to other states.

This article develops the indicated approach, proposing to use a vector database for semantic search of models in the library based on the similarity of load patterns. This idea is applied to solve the key task: to ensure the proactivity of control at the beginning of a new state, when data for training a new model is still insufficient. But it is already possible to assess the similarity of the current situation with previously observed ones, select corresponding models from the library, prepare a predictor, and based on the forecast prepare a control decision.

Method of dynamic local approximation by neural network models with library and vector search of k -nearest models

The proposed development of the DLANNLIB method consists in the addition of a new component – a vector database, which helps implement the mechanism of search and selection of models from the library. The name DLANNLIB k NM (*dynamic local approximation by neural network models with library with k -nearest models vector search*) will be used in the further exposition.

¹NASA-HTTP // Internet Traffic Archive [Electronic resource]. URL: <https://ita.ee.lbl.gov/html/contrib/NASA-HTTP.html> (date of access: 09.09.2025).

Formation of vector representations. For each neural network model, saved in the library, additionally is formed its vector representation (embedding). This feature vector is linked to a specific system state S_i , the load pattern on which the model was trained.

The vector can include various statistical metrics of the training dataset (mean, variance, entropy, correlation dimension, etc.). However, for determining some of them, such as the correlation dimension by the Grassberger – Procaccia method [3] is required a large number of samples and the calculation process itself can take additional time, reducing the operational efficiency.

Most naturally to use as a vector the dataset on which the model was trained or its part. Here various options are possible. The simplest is the use of a vector, built from the first or last samples of the training set, with dimensionality equal to the number of features in the model (window size).

Vector search for a new pattern. Upon the system's transition to a new state S_{i+1} and the arrival of the first utilisation data, before the readiness of a dataset for training a new model, the following actions are performed:

1. Based on these data, the feature vector of the current pattern is determined.
2. A search in the vector database is conducted by the metric of scalar product, cosine similarity, Euclidean, Manhattan distances and other metrics for finding a certain number of the closest vectors by the nearest neighbours method (k -NN).

3. Previously trained neural network models, corresponding to the found vectors, are extracted from the library.

Construction and use of an ensemble predictor. The found models are used to construct an ensemble predictor for the current state. The weights of models in the ensemble can be determined:

- by distance (inversely proportional to the distance between the vector of the current pattern and the vector of the found model);
- quality of the found models (inversely proportional to the errors of the models at the validation stage);
- a hybrid option using both distances and validation errors simultaneously.

The resulting forecast is formed as a weighted sum of the forecasts of individual models of the ensemble.

Training a new model and updating the database. In parallel with the operation of the ensemble predictor, the system continues to accumulate data for the new state. As soon as a sufficient amount of data (samples) has been accumulated, the standard process of training a new local model by the DLANN method is launched. After successful training:

1. The new model together with metadata is saved into the library.
2. For it, a vector representation is determined based on the training sample (sample elements, metadata, statistical parameters).
3. This vector is added to the vector database and linked to the new model.

Forgetting mechanism. In the process of the system's operation, new models and data accumulate in the vector database. A forgetting mechanism can be implemented, which allows limiting the number of semantically close models, leaving only the best models, with minimal validation error.

Model system

Model system consists of a set of computational modules, on which operate instances of application software, between which is carried out balancing of external requests. For request generation is used a load system, allowing to set load and receive statistics on processed requests. The system supports a scaling mechanism (state change). In fig. 2 is presented a diagram of the model system. For managing service configurations is used the Infrastructure as a Code (IaC) approach in which service configurations are described in YAML files.

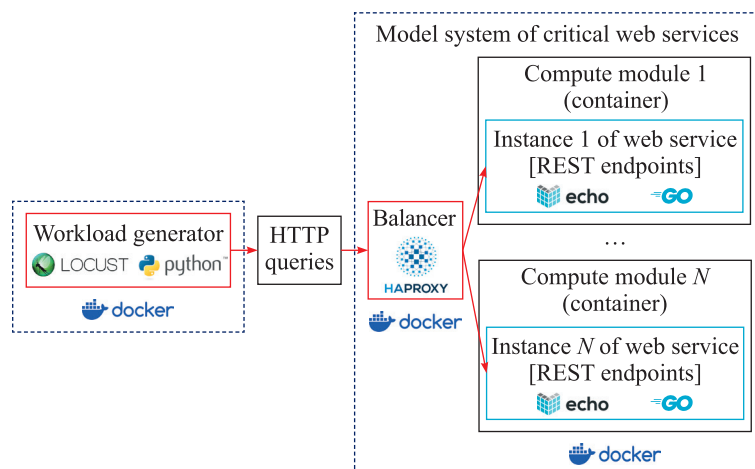


Fig. 2. Diagram of the model system

We will describe the main blocks of the model system.

Block of computational modules. As computational modules is used Docker². This solution allows creating computational modules (Docker containers), managing them, collecting various utilisation metrics. There are ready libraries (e. g., Docker SDK for Python), for working with Docker API Engine. Possible use of other solutions, e. g., Kubernetes clusters, etc.

Block of load balancing. As a load balancer across containers with instances of the application web service is used HAProxy³, which allows automatically adding new instances to the load balancing. Previously in experiments was used load balancing by service name using the mechanism built into Docker, from which had to be abandoned due to uneven balancing when creating new containers.

Block of application web service. As a web application was decided to implement a microservice, based on the popular high-performance minimalist web framework Echo⁴, written in the high-level language Go⁵. Service instances implement functionality in the form of REST API, consisting of several REST endpoints, which process incoming HTTP requests from client systems (external load). Each endpoint is associated with a specific microfunctionality, which has a varying impact on computational resource consumption.

Block of load formation. As a load generator is used Locust⁶, written in Python⁷. Locust has less functionality compared to Apache JMeter⁸, simple to configure, allows describing load profiles in the form of Python classes, less demanding on resources, supports distributed configuration of instances, allows setting the load profile in the form of a function or in the form of pre-prepared data.

Load profiles

For generating synthetic profiles for Locust is written a Python script, which at input receives a description of stages in the form of a list of dictionaries, in which are defined the durations of stages, target functions, width of variance, function for changing variance and others. For generating profiles based on real data is implemented a Python script, which allows forming the structure of stages for launching load for Locust. At the initial stages of system development mainly were used simple synthetic profiles «ramp», «triangle», «trapezoid» and the same profiles with noise. It is worth noting that even simple profiles like «ramp» generate a sufficiently complex dependence of the target metric on time %CPU_Util, for various system states. For example, in fig. 3 and 4 are presented user load profiles «triangle» and «triangle with noise».

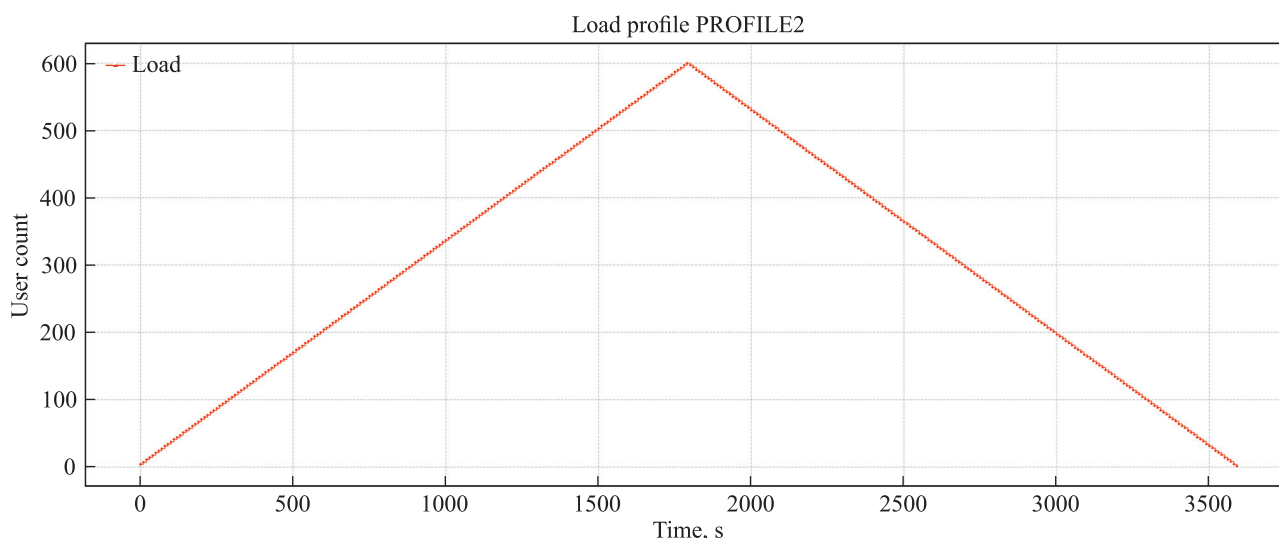


Fig. 3. Load profile «triangle», change the number of users over 1 h

²Docker Compose // Docker Docs : website. URL: <https://docs.docker.com/compose> (date of access: 09.09.2025).

³HAProxy: the reliable, high-performance TCP/HTTP load balancer [Electronic resource]. URL: <https://www.haproxy.org> (date of access: 09.09.2025).

⁴Echo: high-performance, extensible, minimalist Go web framework [Electronic resource]. URL: <https://echo.labstack.com> (date of access: 09.09.2025).

⁵Go: an open-source programming language supported by Google [Electronic resource]. URL: <https://go.dev> (date of access: 09.09.2025).

⁶Locust: an open-source load testing tool [Electronic resource]. URL: <https://locust.io> (date of access: 09.09.2025).

⁷Python: programming language [Electronic resource]. URL: <https://www.python.org> (date of access: 09.09.2025).

⁸Apache JMeter: open-source software [Electronic resource]. URL: <https://jmeter.apache.org> (date of access: 09.09.2025).

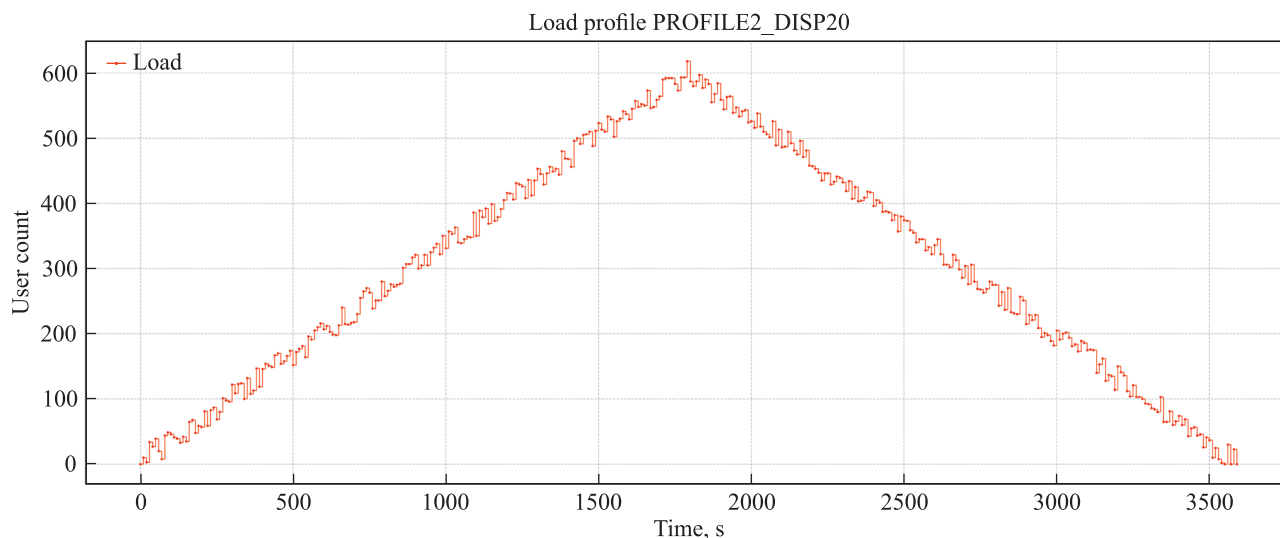


Fig. 4. Load profile «triangle with noise», change the number of users over 1 h

After checking the operation of the control system on synthetic profiles, was chosen a reference load profile, which is often used by other research groups [4]. This profile describes the number of user requests per unit time (*requests per second*, RPS) to various endpoints of the web service of the NASA Kennedy Space Centre in Florida over 24 h from 6 July 1995. The data are presented in public access in the Internet Traffic Archive repository on the website of the Lawrence Berkeley National Laboratory⁹.

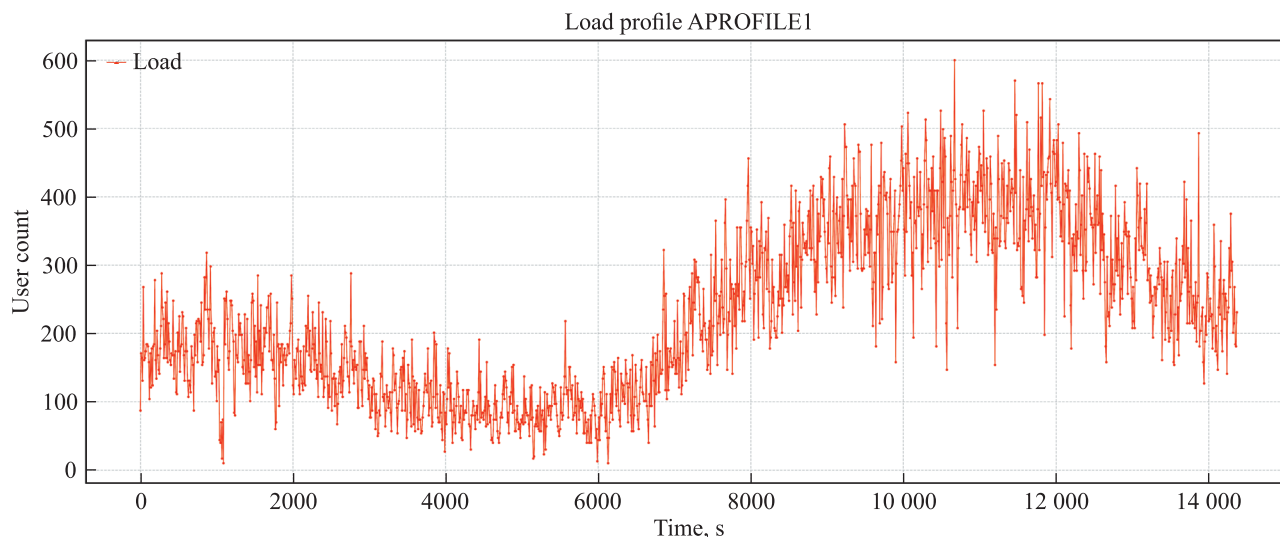


Fig. 5. Load profile «NASA_06.07.1995_4H», change the number of users over 4 h

The original profile was compressed from 24 to 4 h and scaled to 600 users (while the original number of samples was preserved) and converted into a load profile for Locust. In fig. 5 is presented the resulting user load profile «NASA_06.07.1995_4H».

Multi-model automated control system

For managing resources of the model IT system was developed an automated control system (ACS), which in real time receives utilisation data of computational modules and makes a decision on scaling the controlled system. Uses a combination of reactive and proactive control. In this case, for each state of the controlled system automatically is formed its own dataset, is trained its own neural network model, is formed a feature vector of the current pattern based on this dataset, the model and its corresponding vector are saved into the model library, are performed predictions of resource utilisation parameters by different predictors (based on the trained model for the state and model selected from the model library using vector search).

⁹NASA-HTTP // Internet Traffic Archive [Electronic resource]. URL: <https://ita.ee.lbl.gov/html/contrib/NASA-HTTP.html> (date of access: 09.09.2025).

The ACS compares current data of average load across computational modules and forecast results for a specific system state and makes a decision on state change (scaling), depending on the achievement of threshold values of resource utilisation.

Architecturally, the system contains seven main modules, which are presented in fig. 6.

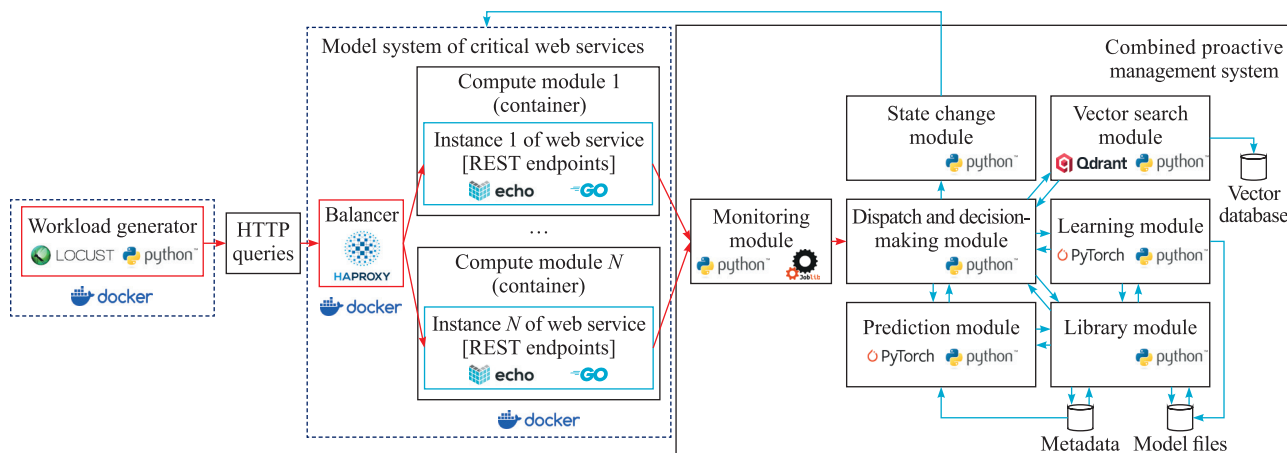


Fig. 6. Diagram of the ACS and the model system

The model IT service, load system, and ACS are described in detail by the authors in the article [2]. As the database is used SQLite¹⁰. The approach with vector search adds new components to the scheme – a vector database and a vector search module. As the vector database is used either a custom lightweight solution in Python, or the `sqlite-vec` extension¹¹, or a ready-made vector database Qdrant¹².

Monitoring module. The monitoring module is responsible for collecting performance metrics of the system's computational modules, as well as adding and removing new metrics (when the system state changes). To correctly collect metrics, the Joblib library¹³ was used, which allows for parallel queries and timely retrieval of metric values for all modules.

State change module. The state change module is responsible for sending control commands (which change the state of the controlled system) and monitoring the correctness of the state change.

Dispatch and decision-making module. The dispatch and decision-making module is the central one. It and the other modules are implemented in Python. At the initialisation stage, resource utilisation thresholds (system SLOs) are set, the achievement of which leads to a change in the state of the controlled system. Thresholds for adding (A) and removing (D) computational resources are set separately. Three control signals are used – the current value of resource utilisation, the forecast based on the model library, and the forecast based on the model trained for the current state. The resulting signal is determined as the maximum of the three signals. An activation mechanism for signals is introduced, which blocks or activates signals in the required sequence. This process is controlled by the parameters `cool_reactive`, `cool_proactive`, `cool_proactive_lib`. This mechanism determines the number of cycles during which no actions to change the state are performed in the system and which signals influence the change of the system state. A parameter is set that determines the number of data accumulation iterations for one state to create a model (M). A lead time (Z) is set – the number of forecast points into the future. In the main process, after the initialisation stage, a cycle is implemented, at each iteration of which the set of the computational modules of the controlled system is updated, the current values of computational module utilisation are obtained, and a decision is made to change the state of the controlled system. Metrics are collected with a period of ~2 s. Historical data on the utilisation of each computational module for different system states are saved in the form of CSV files. In RAM, data on module utilisation for the current state are accumulated in a dictionary. Based on the most recently obtained data, in the main process, the average value across the set of computational modules is calculated. If the maximum of the three active signals exceeds the threshold for addition (A), a decision is made to add resources to the controlled system. If it is less than the removal threshold (D), a decision is made to remove resources. After a decision is made, a command is sent to the state change block. This process implements mechanisms for non-blocking interaction with processes that run in parallel to the main process, responsible for creating and training neural networks and

¹⁰SQLite: C-language library [Electronic resource]. URL: <https://www.sqlite.org> (date of access: 09.09.2025).

¹¹SQLite-Vec: a vector search SQLite extension [Electronic resource]. URL: <https://alegarcia.xyz/sqlite-vec> (date of access: 09.09.2025).

¹²Qdrant: vector search engine [Electronic resource]. URL: <https://qdrant.tech> (date of access: 09.09.2025).

¹³Embarrassingly parallel for loops // Joblib: running Python functions as pipeline jobs [Electronic resource]. URL: <https://joblib.readthedocs.io/en/latest/parallel.html> (date of access: 09.09.2025).

forecasting utilisation parameters for a specific state with a given lead time. This mechanism is implemented using the multiprocessing library¹⁴.

Training and prediction modules. All modules are implemented using Python, using the PyTorch library¹⁵. The training module creates neural network models based on datasets for different states of the controlled system and saves the models and their corresponding vectors into the neural network model library. Each model is encoded by a state index. The prediction module loads models from the neural network model library by index and performs utilisation forecasts with a certain lead time for different states of the controlled system. When forecasting, we assume that each subsequent term of the series depends on a certain number of previous elements of the series, i. e. shifted in time previous values of the original series are used as independent autoregression parameters. The number of such parameters determines the sliding window for the series. We choose the window size (τ_w) equal to 30 elements (determined empirically). We assume that the main contribution to the approximating function for the next element of the series is made by a combination of the 30 previous elements of the series.

A neural network with 1 hidden layer and 1 output neuron is used. The number of neurons in the hidden layer (90) is 3 times greater than the temporal window size. In the output layer, there is 1 neuron, as a sum element. A fully connected linear layer `nn.Linear` is used. The activation function is ReLU. For regularisation, dropout of part of the neurons is used (dropout with a neuron dropout probability of 0.015). The loss function used is `nn.MSELoss()` (mean squared error). The optimisation method is `torch.optim.Adam`, learning rate 0.002. Preprocessing of the original series with scaling to the interval $[0, 1]$ is not performed, as it introduces additional error into the original data and leads to additional overhead for scaling before and after training (we utilise the property of the original data, which are bounded from below and above). A short time series is used – 60 samples. The interval between samples is 2 s. The duration of the series is 128 s. The original series is divided into two datasets – train (42 samples, or 70 %) and test (18 samples, or 30 %). Considering that the window size (τ_w) is 30, the number of examples for training is 13. A total of 100 training epochs are used. The batch size is 1. The training time for a neural network with the specified architecture for a dataset is ~ 2 s. The speed of performing a prediction with a lead time of 40 samples is $\ll 1$ s.

Estimation of management quality

To evaluate the quality of management, we consider a problem, where the target function reflects a trade-off between optimising computational resource costs and meeting SLO requirements. The problem variables: $\Omega(t)$ – dynamically changing external system load with uncertainty; $R_k(t)$ – the amount of computational resources allocated to the system at the k^{th} step (system resources change discretely, we assume that the k^{th} step corresponds to state S_k); $\text{Cost}(R_k(t))$ – the cost of resource usage at the k^{th} step; $\text{SLO}_k(t)$ – the SLO metric at the k^{th} step; $\text{Violation}(\text{SLO}_k(t))$ – the penalty for SLO violation. $\text{Cost}(R_k(t))$ and $\text{Violation}(\text{SLO}_k(t))$ are normalised in the interval $[0, 1]$ over all states S_k . The target function to be minimised:

$$\arg \min_A \left(\lambda \sum_{k=1}^N \text{Cost}(R_k[A(\Omega(t))]) + (1-\lambda) \sum_{k=1}^N \text{Violation}(\text{SLO}_k[A(\Omega(t))]) \right)$$

with the constraint on the rate of dynamic resource changes $|R_k - R_{k-1}| \leq \gamma$, where γ is the maximum allowable resource change.

As a result, the problem reduces to minimising the target function, achieved by selecting the optimal algorithm A , which generates N system states under external load $\Omega(t)$ with uncertainty. An exact solution to this problem cannot be constructed, so we propose a heuristic algorithm that estimates the target function. From the set of algorithms, we select the one that minimises the target function.

Experimental environment

The model system, load balancing service and load testing system operate in Docker containers on a virtual machine running a Linux-based OS – Ubuntu 22.04.1 LTS. The control system is deployed directly in the OS. Tests were conducted on various physical hardware. Final tests were performed on a virtual machine with the following parameters: vCPU – 8 cores, RAM – 24 GB.

We evaluate the performance quality of the combined automatic scaling control system using the previously described methodology. We use the quality metric F with an adjustable weight λ to specify the desired

¹⁴Multiprocessing – Process-based parallelism // The Python Standard Library [Electronic resource]. URL: <https://docs.python.org/3/library/multiprocessing.html> (date of access: 09.09.2025).

¹⁵PyTorch : website. URL: <https://pytorch.org> (date of access: 09.09.2025).

cost-to-penalty ratio. A threshold value of 24 ms is used for the response time penalty function. For calculating maximum cost, we use $C_{\max} = 6$.

In experiments, three load profiles are used, in table 1 is presented a brief description.

Table 1

Description of load profiles		
Profile name	Description	Time, h
PROFILE2	Profile «triangle» without dispersion. Linear increasing load from 1 to 600 users over 30 min and linear decreasing load from 600 to 0 users over 30 min	1
PROFILE2_DISP20	Profile «triangle» with random noise. Linear increasing load with random noise from 1 to 600 users over 30 min and linear decreasing load with random noise from 600 to 0 users over 30 min	1
APROFILE1	Profile «NASA_06.07.1995_4H». Based on real load	4

Control algorithms

Each profile was tested on three different algorithms. In the algorithm name are encoded the main operating parameters. Decoding of the code is presented in table 2. In each test group, corresponding to a specific profile, the test with the reactive algorithm is considered as the baseline. The operation of other algorithms is compared by quality metrics with the baseline algorithm.

Table 2

Decoding of algorithm codes			
Algorithm code	Type	Main operating parameters	Value
A90.50.R60	Reactive	Addition threshold, %CPU_Util	90
		Removal threshold, %CPU_Util	50
		Reactive component activation period, samples	60
A90.50.R60P70F	Hybrid (reactive + DLANN)	Addition threshold, %CPU_Util	90
		Removal threshold, %CPU_Util	50
		Reactive component activation period, samples	60
		Proactive component activation period, samples	70
A90.50.R60P70L60_70FNNL_VEC1	Hybrid (reactive + DLANLIB_kNM)	Addition threshold, %CPU_Util	90
		Removal threshold, %CPU_Util	50
		Reactive component activation period, samples	60
		Proactive component activation period, samples	70
		Proactive component activation period (library), samples	60

Let us briefly describe the activation mechanism of control signals for the algorithm with library and vector model search A90.50.R60P70L60_70FNNL_VEC1. A stabilisation period equal to 60 samples (120 s) is used for forced model formation for each state (also possible is the variant used in earlier versions of the ACS, where a model is not formed for each state). After the end of this period, in the decision-making mechanism, signals related to the reactive part (current values of the collected metric) and the forecast based on the model library are activated, and after 10 samples (20 s), the signal related to the forecast using the model trained for the current state is activated, and simultaneously the signal related to the forecast based on the model library is deactivated (we assume that the model trained on data of the current state has greater expertise within this state). Schematically, the activation mechanisms for the algorithms specified in table 2 are shown in fig. 7.

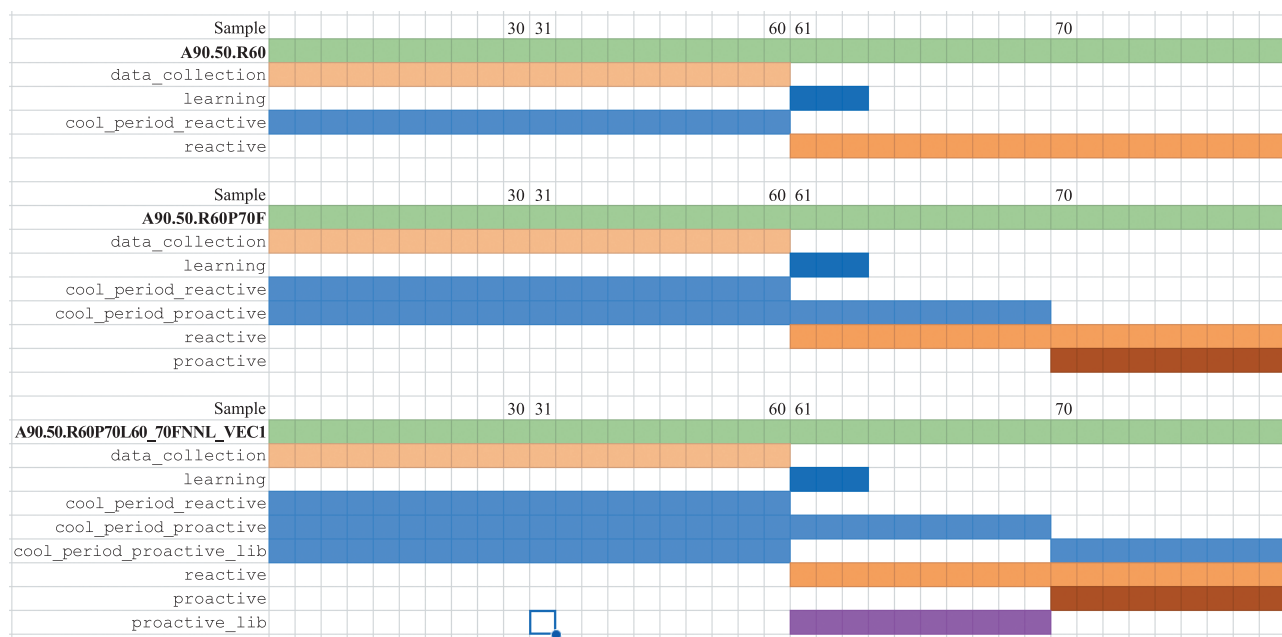


Fig. 7. Diagram of control signal activation

In table 3 are presented the results of experiments on the operation of the ACS for three different load profiles. Quality metrics were calculated after conducting the experiments, based on the obtained statistics of the load system and logs of the control system operation.

Table 3

Results of experiments

Profile	Algorithm code	Normalised values		Target function		
		C^{norm}	V^{norm}	$2F_{1/2}$	$F_{2/3}$	$F_{1/3}$
PROFILE2	A90.50.R60	0.367	0.710 (+204.6 %)	1.077 (+34.6 %)	0.482	0.596
	A90.50.R60P70F	0.403 (+9.7 %)	0.596 (+155.8 %)	1.000 (+24.9 %)	0.468	0.532
	A90.50.R60P70L60_70FNNL_VEC1	0.567 (+54.4 %)	0.233	0.800	0.456	0.345
PROFILE2_DISP20	A90.50.R60	0.341	0.705 (+51.8 %)	1.046 (+26.5 %)	0.463	0.584
	A90.50.R60P70F	0.398 (+16.6 %)	0.595 (+28.2 %)	0.993 (+20.1 %)	0.464	0.530
	A90.50.R60P70L60_70FNNL_VEC1	0.363 (+6.3 %)	0.464	0.827	0.397	0.431
APROFILE1	A90.50.R60	0.333	0.743 (+56.6 %)	1.076 (+14.4 %)	0.470	0.606
	A90.50.R60P70F	0.354 (+6.3 %)	0.673 (+42.0 %)	1.027 (+9.3 %)	0.460	0.567
	A90.50.R60P70L60_70FNNL_VEC1	0.466 (+39.6 %)	0.474	0.940	0.469	0.471

Note. Boldface indicates the best (minimum) normalised value for each profile across the group of algorithms. The values shown in the table represent resource cost, SLO violation penalty, and their sum, respectively. Values in parentheses show the percentage increase relative to the best (minimum) value within the same group for the corresponding profile.

Figures 8–13 are used as an illustration of the differences in algorithm operation for load profiles PROFILE2, PROFILE2_DISP20 and APROFILE1. On the graph of average CPU load across computational modules, data for different states are shown in colour; on the lower graph, the system trajectory (dependence of the number of computational modules on time) with state numbers is shown.

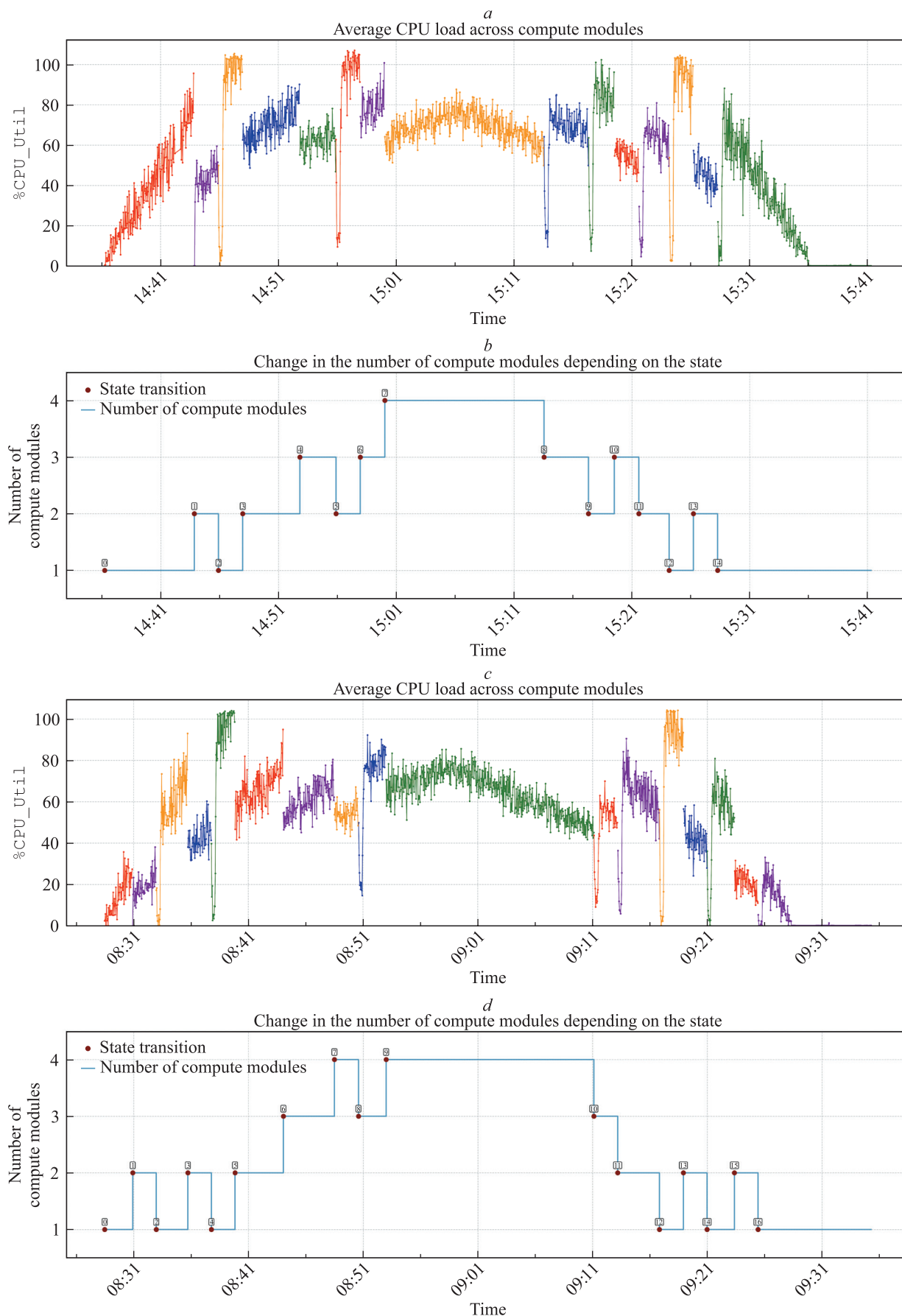


Fig. 8. Illustration of the operation of algorithms A90.50.R60 (*a, b*) and A90.50.R60P70F (*c, d*) for profile PROFILE2

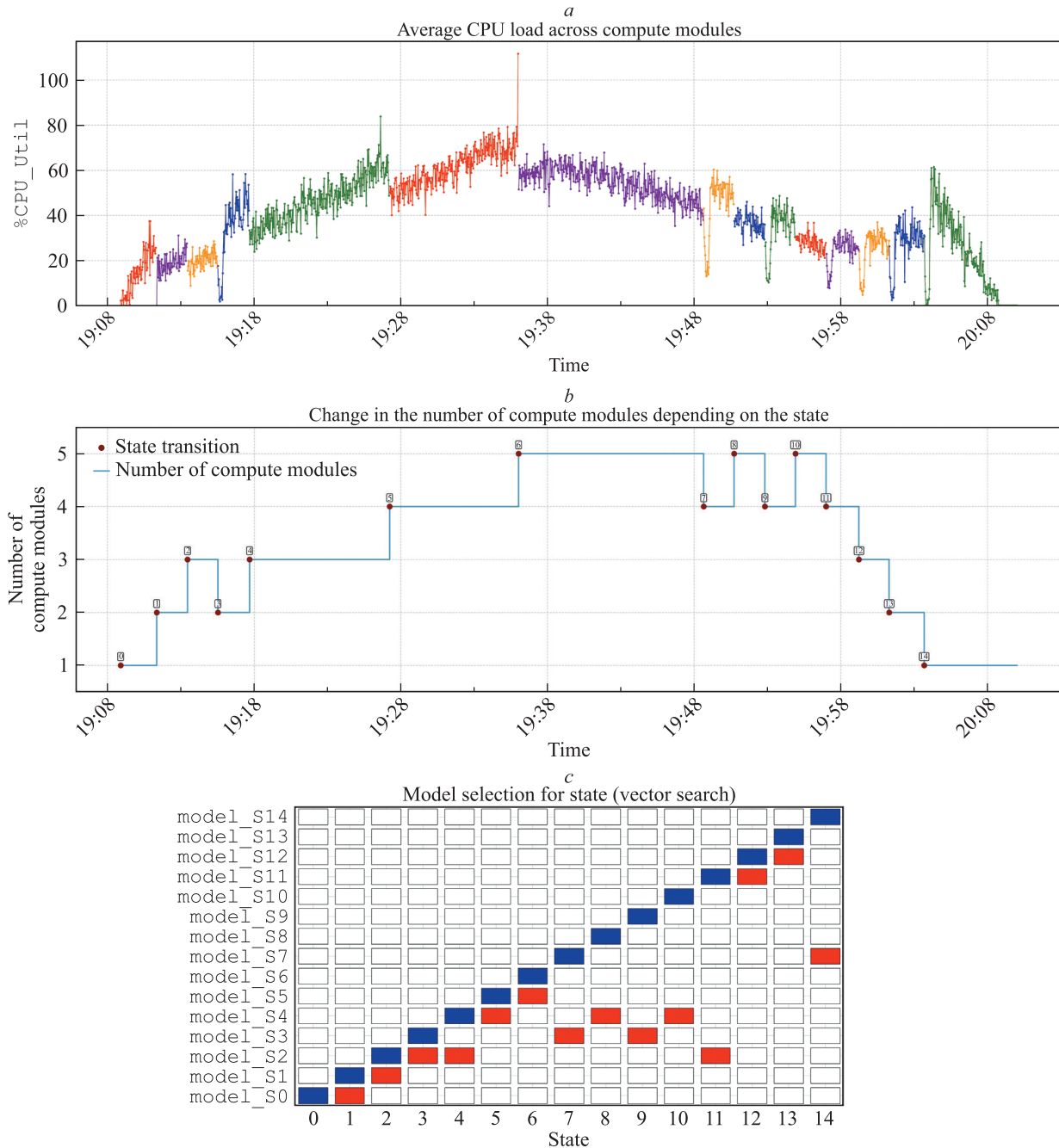


Fig. 9. Illustration of the operation of algorithm A90.50.R60P70L60_70FNNL_VEC1 (*a*, *b*) and model selection using vector search (*c*) for profile PROFILE2

Analysis of the results for profile PROFILE2 shows that in terms of the sum of normalised cost and penalty metric $2F_{1/2}$, the best was the algorithm with library and vector model search A90.50.R60P70L60_70FNNL_VEC1. It is approximately 34.6 % better than the reactive algorithm A90.50.R60 and 24.9 % better than the hybrid algorithm A90.50.R60P70F. If comparing by individual metrics, the most resource-efficient, as expected, is the reactive algorithm A90.50.R60, which is 9.7 % more economical than algorithm A90.50.R60P70F and 54.4 % more economical than algorithm A90.50.R60P70L60_70FNNL_VEC1. In terms of penalties for response time, the best was algorithm A90.50.R60P70L60_70FNNL_VEC1, which is almost 204.6 % better than the reactive algorithm A90.50.R60 and 155.8 % better than hybrid algorithm A90.50.R60P70F. For the target function $F_{1/3}$ with a greater weight towards penalties, the best result was shown by the algorithm A90.50.R60P70L60_70FNNL_VEC1, followed by the algorithm A90.50.R60P70F, which were 73.0 and 54.4 % better than the algorithm A90.50.R60. Figure 9, *c*, shows which models were selected from the library using the DLANNLIB_kNM method for various states under load profile PROFILE2. It is interesting to note that up to state 12, various models trained on increasing load were selected. At the beginning, predominantly the model trained for the previous state was chosen. And only starting from state 12, models trained during decreasing load began to be used.

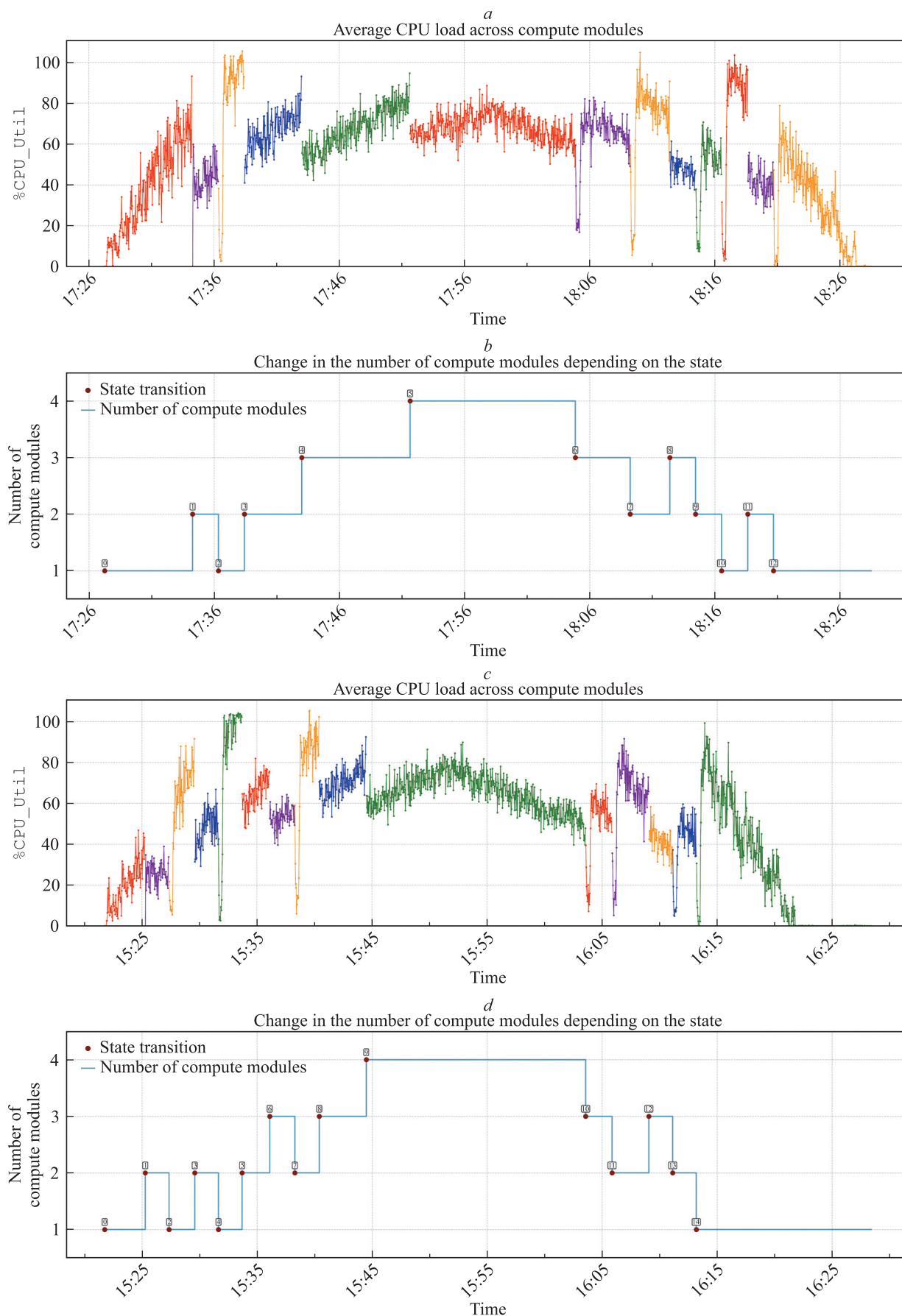


Fig. 10. Illustration of the operation of algorithms A90.50.R60 (*a*, *b*) and A90.50.R60P70F (*c*, *d*) for profile PROFILE2_DISP20

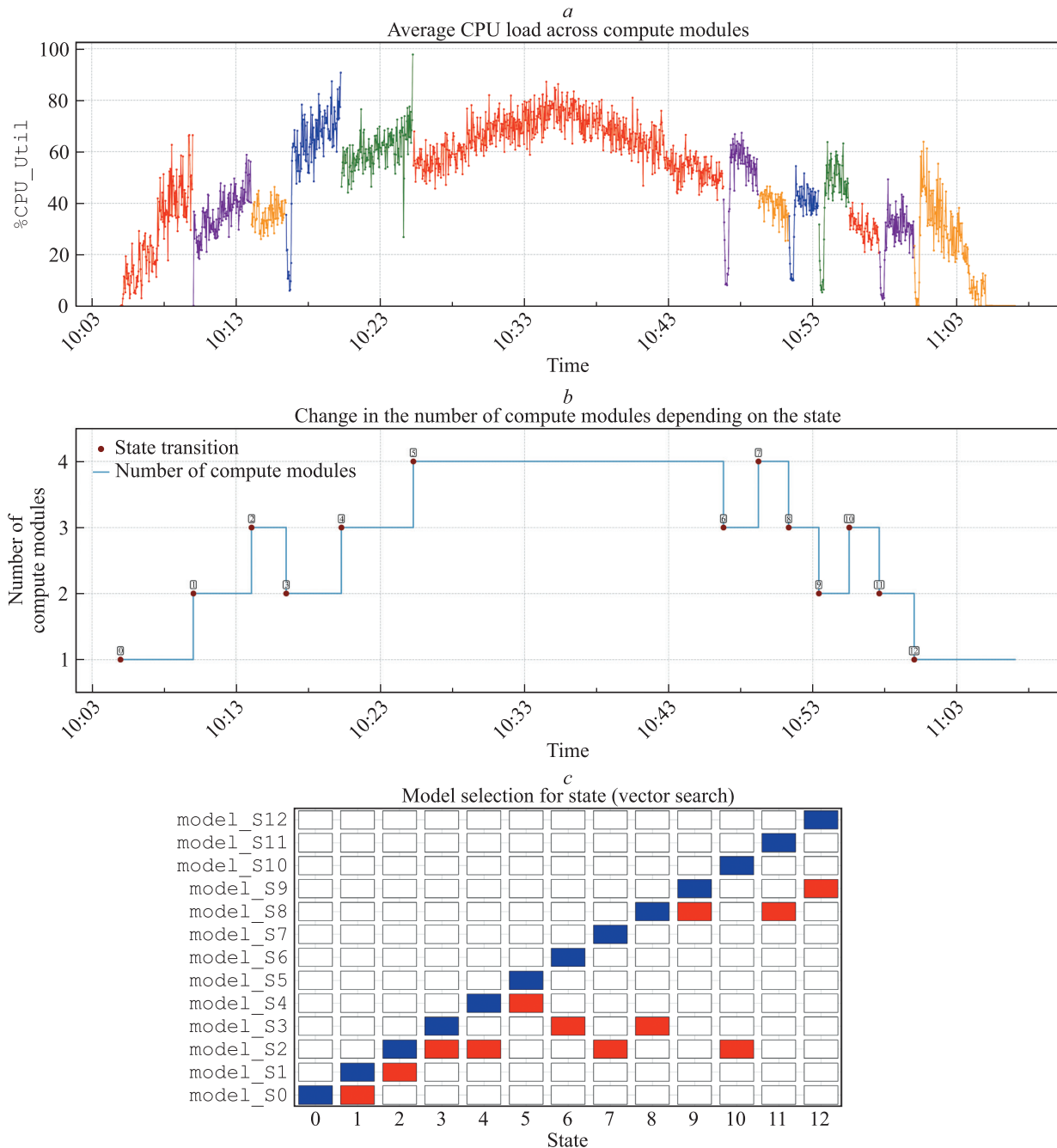


Fig. 11. Illustration of the operation of algorithm A90.50.R60P70L60_70FNNL_VEC1 (*a*, *b*) and model selection using vector search (*c*) for profile PROFILE2_DISP20

Analysis of the results for the more complex profile PROFILE2_DISP20 shows that in terms of the target function $2F_{1/2}$, the best results were again shown by the algorithm A90.50.R60P70L60_70FNNL_VEC1. It is approximately 26.5 % better than the reactive algorithm A90.50.R60 and 20.1 % better than the hybrid algorithm A90.50.R60P70F. If comparing by individual metrics, the most resource-efficient, as expected, is the reactive algorithm A90.50.R60, which is 6.3 % more economical than algorithm A90.50.R60P70L60_70FNNL_VEC1 and 16.6 % more economical than algorithm A90.50.R60P70F. In terms of penalties for response time, the best was algorithm A90.50.R60P70L60_70FNNL_VEC1, which is almost 51.8 % better than the algorithm A90.50.R60 and 28.2 % better than the algorithm A90.50.R60P70F. For the target function $F_{1/3}$ with a greater weight towards penalties, the best result was shown by the algorithm A90.50.R60P70L60_70FNNL_VEC1, followed by the algorithm A90.50.R60P70F, whose results were 35.6 and 23.0 % better than the algorithm A90.50.R60. Figure 11, *c*, shows which models were selected from the library using the DLANNLIB_kNM method for various states under load profile PROFILE2_DISP20. It is interesting to note that up to state 6, various models trained on increasing load were selected. Further, during decreasing load, for states 6, 7 and 8, models trained during increasing load were also selected, but for state 9, a model trained during decreasing load was already chosen.

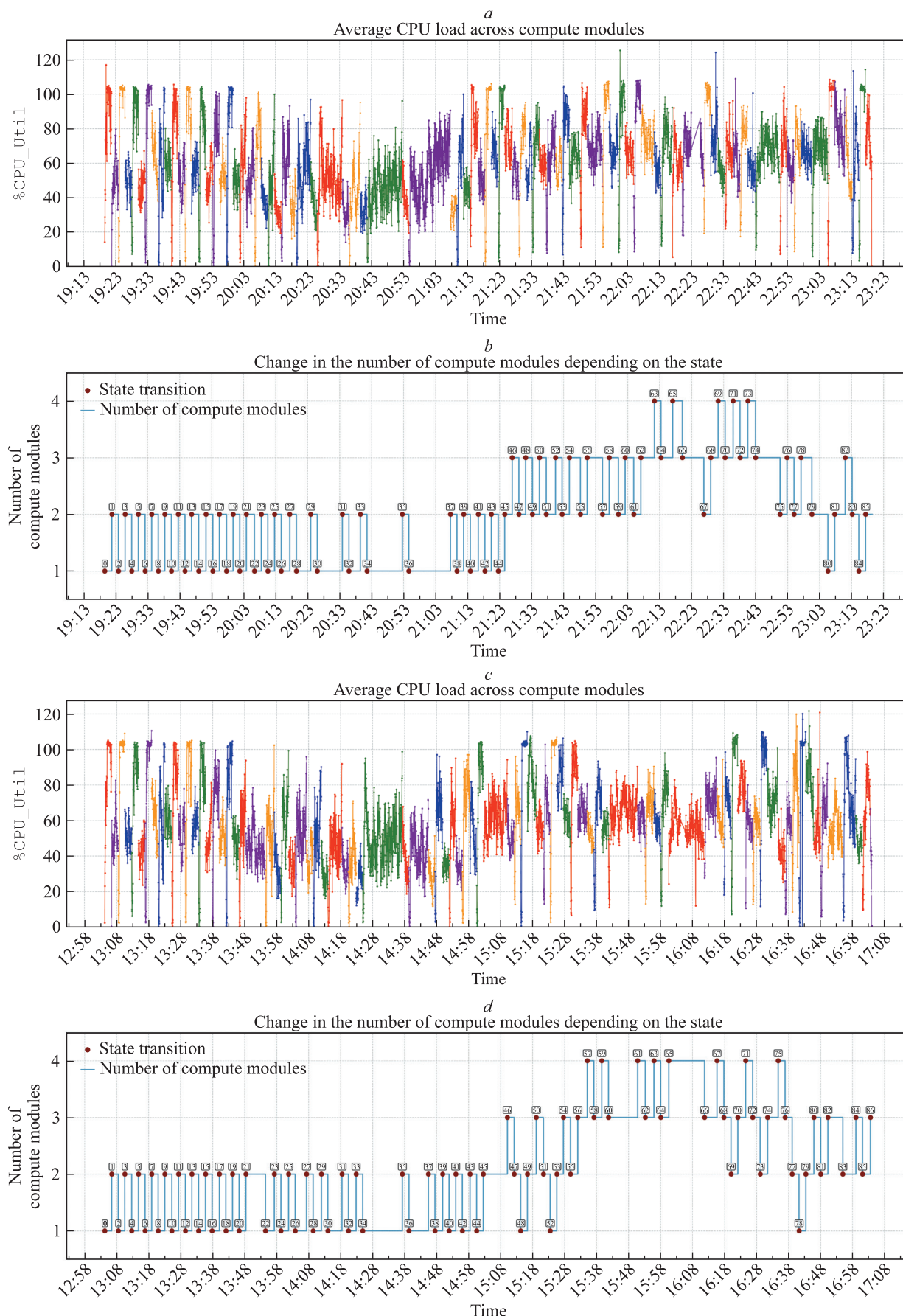


Fig. 12. Illustration of the operation of algorithms A90.50.R60 (*a*, *b*) and A90.50.R60P70F (*c*, *d*) for profile APROFILE1

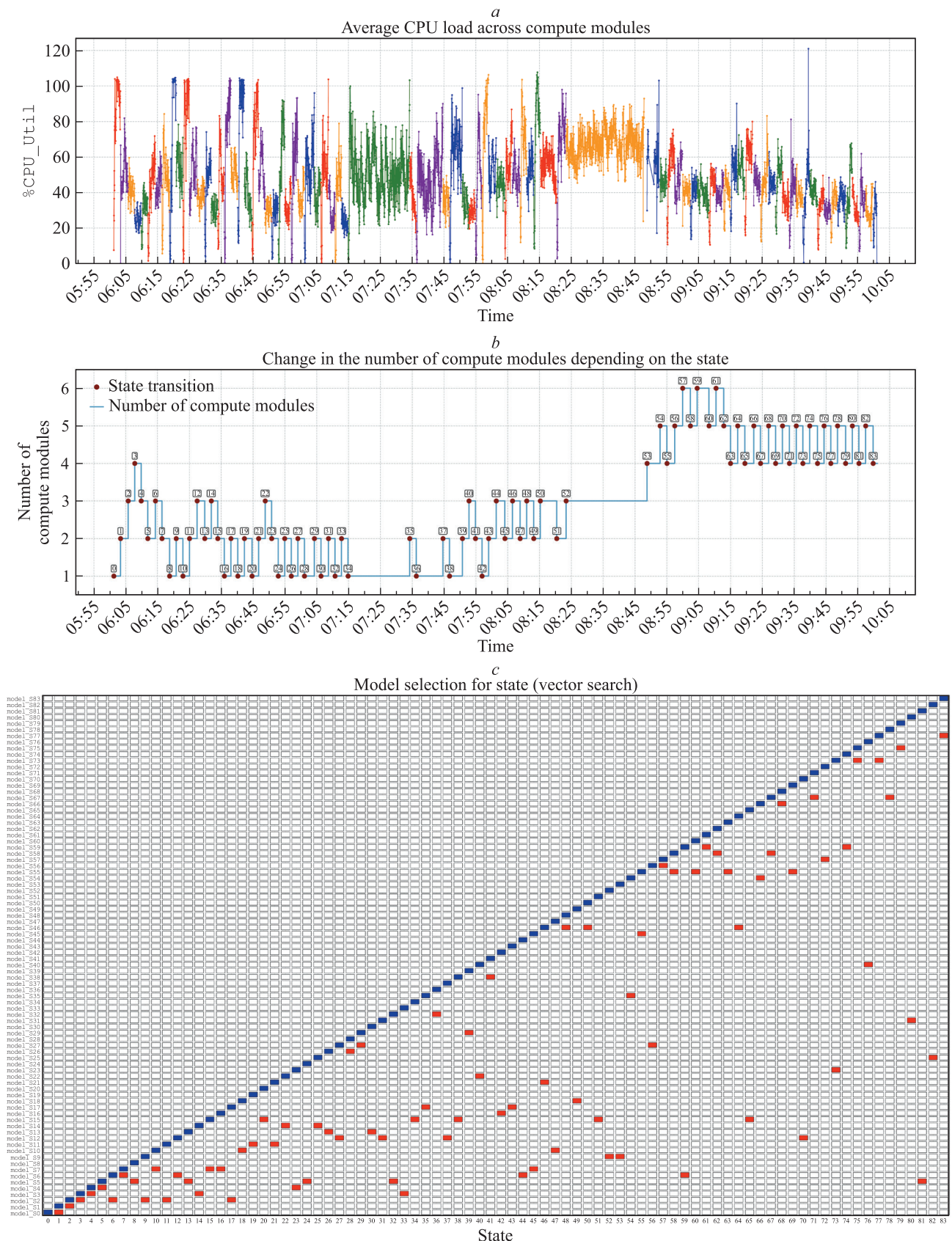


Fig. 13. Illustration of the operation of algorithm A90.50.R60P70L60_70FNNL_VEC1 (*a*, *b*) and model selection using vector search (*c*) for profile APROFILE1

Let us consider the results for the most complex profile APROFILE1. Analysis shows that in terms of the target function $2F_{1/2}$, the best result was shown by the algorithm A90.50.R60P70L60_70FNNL_VEC1. It is approximately 14.4 % better than the reactive algorithm A90.50.R60 and 9.3 % better than the hybrid algorithm A90.50.R60P70F. If comparing by individual metrics, the most resource-efficient is the reactive algorithm A90.50.R60, which is 6.3 % more economical than algorithm A90.50.R60P70F and 39.9 % more economical than algorithm A90.50.R60P70L60_70FNNL_VEC1. In terms of penalties for response time, algorithm A90.50.R60P70L60_70FNNL_VEC1 is the best, 56.6 % better than the reactive algorithm A90.50.R60 and 42.0 % better than the hybrid algorithm A90.50.R60P70F. For the target function $F_{1/3}$ (greater weight towards penalties), the best result is with algorithm A90.50.R60P70L60_70FNNL_VEC1, followed by the algorithm A90.50.R60P70F, whose results were 28.6 and 20.2 % better than the algorithm A90.50.R60. Figure 13, c, shows which models were selected from the library using the DLANNLIB_kNM method for various states under load profile APROFILE1. At the beginning, when the library does not contain many models, models for the previous state are predominantly chosen; with further filling of the library, the choice becomes different. It is interesting to note that in earlier studies for the most complex profile APROFILE1, the DLANNLIB_LM method (the simplest variant, where the model for the previous state was selected from the library) was only 5 % better than the reactive one.

The obtained results show that, as in previous studies, algorithms that work well on synthetic data seriously reduce their effectiveness on a complexly structured profile based on real load, but at the same time, new ideas related to complicating the predictor based on using vector search over a model library allow improving the result, even when using the simplest version of the algorithm (where only one model from the library is used to build the predictor, not k -nearest models).

Conclusions

The work proposes a new method, increasing the operational efficiency of the control system. The approach implements adaptation to new load patterns, associated with uncertainty, through approximation of data in real time using neural network models. The obtained models are saved into a library together with vector descriptions of load patterns, which allows subsequently promptly selecting suitable models for new system states using vector search by semantic similarity and building an ensemble predictor.

The results of experiments show that algorithms that work well on synthetic loads seriously reduce their effectiveness on a complexly structured profile based on real load. But using a predictor built with vector search allows improving the result, even when using the simplest version of the algorithm.

In terms of the aggregate of indicators when working with irregular and complex load, hybrid algorithms turned out to be more effective than reactive ones. At the same time, the best result was shown by the algorithm using vector search.

References

1. Straesser M, Grohmann J, von Kistowski J, Eismann S, Bauer A, Kounev S. Why is it not solved yet? Challenges for production-ready autoscaling. In: Association for Computing Machinery. *ICPE'22. Proceedings of the 2022 ACM/SPEC International conference on performance engineering; 2022 April 9–13; Beijing, China*. New York: Association for Computing Machinery; 2022. p. 105–115. DOI: 10.1145/3489525.3511680.
2. Krasnoprosin V, Starovoitov A. Real-time management of critical IT service using a multi-model approach. In: Kamil Aida-zade, editor. *2025 6th International conference on problems of cybernetics and informatics (PCI); 2025 August 26–28; Baku, Azerbaijan*. [S. l.]: Institute of Electrical and Electronics Engineers; 2025. p. 1–5. DOI: 10.1109/PCI66488.2025.11219764.
3. Grassberger P, Procaccia I. Measuring the strangeness of strange attractors. *Physica D: Nonlinear Phenomena*. 1983;9(1–2): 189–208. DOI: 10.1016/0167-2789(83)90298-1.
4. Aslanpour MS, Ghobaei-Arani M, Toosi AN. Auto-scaling web applications in clouds: a cost-aware approach. *Journal of Network and Computer Applications*. 2017;95:26–41. DOI: 10.1016/j.jnca.2017.07.012.

Received 24.12.2025 / revised 19.01.2026 / accepted 29.01.2026.

ВЫДЕЛЕНИЕ ОБЪЕКТОВ ИЗОБРАЖЕНИЙ ДЛЯ АВТОНОМНОЙ ЛОКАЛИЗАЦИИ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ И ПОСТРОЕНИЯ КАРТЫ ОКРУЖАЮЩЕГО ПРОСТРАНСТВА

В. Ю. ЦВЕТКОВ¹⁾, В. В. РАБЦЕВИЧ²⁾, МА ЦЗЮНЬ¹⁾, М. И. ЗОРЬКО¹⁾

¹⁾Белорусский государственный университет информатики и радиоэлектроники,
ул. П. Бровки, 6, 220013, г. Минск, Беларусь

²⁾Стормнэт тэкнолоджис, пр. Независимости, 168, корп. 2, 220141, г. Минск, Беларусь

Аннотация. Рассмотрена задача автономной локализации беспилотного летательного аппарата и построения карты его окружения. Показана низкая устойчивость детекторов ключевых точек к изменению погодных и временных условий формирования изображений окружающего пространства по сравнению с устойчивостью детекторов линий. Для повышения устойчивости локализации беспилотных летательных аппаратов в качестве ключевых объектов идентификации предложено использовать однородные по яркости области изображений (сегменты). Разработан алгоритм сегментации по локальным минимумам градиента яркости полутоновых изображений, получаемых из исходных цветных изображений окружающего пространства беспилотного летательного аппарата. Для выравнивания скоростей роста областей локальных минимумов предложен алгоритм волновой сегментации с автоматическим остановом, основанный на определении новых начальных и дополнительных точек роста областей с помощью монотонно изменяющегося порога, а также на учете скорости изменения градиента яркости вдоль траектории роста области в критерии останова. Показано, что предложенные алгоритмы обеспечивают более устойчивую локализацию областей при изменении погодных и временных условий формирования изображений по сравнению с известными алгоритмами выделения линий и ключевых точек.

Ключевые слова: одновременная локализация и картографирование; волновая сегментация изображений; выравнивание областей; идентификация объектов изображений; детекторы точек; детекторы линий.

Благодарность. Работа выполнена при финансовой поддержке Белорусского республиканского фонда фундаментальных исследований и Национального фонда естественных наук Китая (проект T25KI-054 «Технология многокамерной локализации и построения карты открытого пространства»).

Образец цитирования:

Цветков ВЮ, Рабцевич ВВ, Ма Цзюнь, Зорько МИ. Выделение объектов изображений для автономной локализации беспилотных летательных аппаратов и построения карты окружающего пространства. *Журнал Белорусского государственного университета. Математика. Информатика.* 2026; 1:156–162 (на англ.).
EDN: OYZDOV

For citation:

Tsviatkou VYu, Rabtsevich VV, Ma Jun, Zorko MI. Detection of image objects for autonomous localisation of unmanned aerial vehicles and map formation of the surrounding space. *Journal of the Belarusian State University. Mathematics and Informatics.* 2026;1:156–162.
EDN: OYZDOV

Авторы:

Виктор Юрьевич Цветков – доктор технических наук, профессор; заведующий кафедрой инфокоммуникационных технологий факультета информационной безопасности.
Виолетта Викторовна Рабцевич – преподаватель.
Ма Цзюнь – ассистент кафедры инфокоммуникационных технологий факультета информационной безопасности.
Михаил Иванович Зорько – ассистент кафедры инфокоммуникационных технологий факультета информационной безопасности.

Authors:

Viktor Yu. Tsviatkou, doctor of science (engineering), full professor; head of the department of infocommunication technologies, faculty of information security.
vtsvet@bsuir.by
Viyaleta V. Rabtsevich, lecturer.
rabtsevichv@gmail.com
Ma Jun, assistant at the department of infocommunication technologies, faculty of information security.
majun@bsuir.by
Mikhail I. Zorko, assistant at the department of infocommunication technologies, faculty of information security.
m.zorko@bsuir.by

DETECTION OF IMAGE OBJECTS FOR AUTONOMOUS LOCALISATION OF UNMANNED AERIAL VEHICLES AND MAP FORMATION OF THE SURROUNDING SPACE

V. Yu. TSVIATKOU^a, V. V. RABTSEVICH^b, MA JUN^a, M. I. ZORKO^a

^aBelarusian State University of Informatics and Radioelectronics, 6 P. Browki Street, Minsk 220013, Belarus

^bStormnet Technologies, 168 Niezaliezhnasci Avenue, 2 building, Minsk 220141, Belarus

Corresponding author: V. Yu. Tsviatkou (vtsvet@bsuir.by)

Abstract. The problem of autonomous localisation of an unmanned aerial vehicle and construction of a map of its environment is considered. The low stability of keypoint detectors to changing weather and time conditions of image formation in the surrounding space is demonstrated compared to the stability of line detectors. To improve the stability of unmanned aerial vehicles localisation it is proposed to use image regions (segments) with uniform brightness as key identification objects. An algorithm for segmentation based on local minima of the brightness gradient of halftone images obtained from the original colour images of the unmanned aerial vehicle's surrounding space has been developed. To equalise the growth rates of local minima a wave segmentation algorithm with automatic stopping is proposed. This algorithm is based on identifying new initial and additional growth points of regions using a monotonically varying threshold and taking into account the rate of change of the brightness gradient along the region growth trajectory in the stopping criterion. It is shown that the proposed algorithms provide more stable localisation of areas, when changing weather and time conditions for image formation in comparison with known algorithms for extracting lines and key points.

Keywords: simultaneous localisation and mapping; SLAM; wave segmentation of images; region growing; image object identification; point detectors; line detectors.

Acknowledgements. This work was carried out with the financial support of the Belarusian Republican Foundation for Fundamental Research and the National Natural Science Foundation of China (project T25KI-054 «Outdoor multi-camera simultaneous localisation and mapping technology»).

Introduction

Simultaneous localisation and mapping (SLAM) technology is widely used in autonomous mobile robot systems, allowing for the construction of a map of the environment and spatial orientation [1–6]. This technology is based on the comparison of images generated at different points in time during the mobile robot's movement, the binding of objects identified in the images to a digital map of the area, and the identification of objects in the images and the map. A pressing challenge is to ensure the operation of such systems in open space conditions with significant changes in weather and time conditions for image recording.

The aim of the work is to increase the stability of localisation of key objects under changing weather and time conditions for image formation.

Statement of the problem

Basic SLAM methods utilise low-level geometric objects – points extracted by analysing the brightness gradient of images – for identification [1–6]. Newer SLAM methods [7–9] typically utilise semantic objects, whose intrinsic identification parameters are less dependent on weather and time conditions during image formation. However, detecting semantic objects requires significant computational effort. Therefore, relatively simple detectors that extract less complex, high-level geometric objects – lines and segments – are of interest. These objects exhibit reduced localisation stability in response to changing weather and time conditions, while the computational complexity of detection is reduced compared to semantic segmentation.

Segmentation by local minima of the brightness gradient of grayscale images

To identify areas of uniform brightness that are resilient to changes in weather and time conditions during image formation we propose a segmentation algorithm based on local minima of the brightness gradient, MinLGS (minimum local gradient segmentation). This algorithm is based on growing neighbourhoods of local minima of the brightness gradients of grayscale images obtained from the original colour images of the unmanned aerial

vehicles (UAV) surroundings using standard colour transformations. This algorithm differs from known segmentation algorithms by taking into account the maximum brightness difference in the neighbourhood of a 3×3 pixel as the gradient value for thresholding the smoothest image areas; by identifying not only point-based but also the most stable area-based local minima as the starting point for segmentation; by ensuring consistency in the growth rates of the local gradient minima regions. The algorithm consists of the following steps:

- 1) converting a colour image $I_{\text{RGB}} = \|i_{\text{RGB}}(y, x)\|_{(y=\overline{0, Y-1}, x=\overline{0, X-1})}$ of $Y \times X$ pixels into a grayscale image $I_{\text{HT}} = \|i_{\text{HT}}(y, x)\|_{(y=\overline{0, Y-1}, x=\overline{0, X-1})}$;
- 2) formation of a gradient image $I_G = \|i_G(y, x)\|_{(y=\overline{0, Y-1}, x=\overline{0, X-1})}$, whose pixel values $i_G(y, x)$ are calculated using the expression $i_G(y, x) = \max(|i_{\text{HT}}(y+m, x+n) - i_{\text{HT}}(y, x)|), m = \overline{-1, 1}, n = \overline{-1, 1}$;
- 3) cycle through brightness levels $q = \overline{0, 255}$. This step includes search for local minima of gradients in the gradient image I_G at level q using the algorithm proposed in work [10], which identifies not only point but also stable area local minima, and coordinated growth of regions of local gradient minima by joining existing regions of pixels, whose brightness gradient value corresponds to the q level;
- 4) merging adjacent segments with close brightness gradient boundary values. The merged segment inherits the smallest number among the merged segments;
- 5) reassigning numbers to segments to ensure continuity and reduce the maximum number row value.

Gradient-wave segmentation of halftone images

In order to improve the accuracy of division into regions of halftone images, containing objects on a uniform background, and to simplify the processing by eliminating the stage of preliminary determination of the initial growth points a Gradient-Wave Region Growing (GWRG) algorithm is proposed. This algorithm is based on wave growing of regions of local maxima with their selection in the order of decreasing values according to a monotonically changing threshold and analysis of the intensity gradient in the direction of region growth. The algorithm differs from known segmentation algorithms in the following way:

- 1) defining as initial growth points of regions single or groups of adjacent pixels, whose intensities coincide with the value of a threshold that monotonically changes from the maximum value to the minimum and determines the iteration of processing new pixels selected for joining to existing regions or forming new regions;
- 2) matching the growth rates of regions (joining new pixels to regions) by taking into account the values of the first derivative with respect to intensity;
- 3) cessation of region growth upon a decrease relative to the threshold of the local value of the second derivative with respect to intensity, calculated along the growth chain of the region connecting adjacent pixels, namely the base pixel located inside the region, the boundary pixel of the region joined to it at the previous iteration and the candidate pixel processed at the current iteration for joining to the boundary pixel of the region.

The algorithm is built on a cycle through brightness levels q ($q = \overline{2^B - 1, 0}$) in which the following is performed:

- 1) determination of significant elements in a matrix of $M_P = \|m_P(y, x)\|_{(y=\overline{0, Y-1}, x=\overline{0, X-1})}$ pixels ($m_P(y, x) \in [0, 2^B - 1]$, where B is the pixel bit depth) with the size of the $Y \times X$ relative to the threshold q (function F_{MP});
- 2) growing regions (function F_{RG}) is the formation of a segmentation matrix $M_S = \|m_S(y, x)\|_{(y=\overline{0, Y-1}, x=\overline{0, X-1})}$, the values of whose elements indicate the numbers of segments n_S to which they belong ($n_S \in [1, N_S]$, where N_S is the number of segments);
- 3) selection of growth grains of regions – single-pixel and multi-pixel (areal) local maxima (function F_{SP}) – pixels that were not connected to any region at the threshold value q ;
- 4) initial growth of the region (function F_{PRG}).

The function F_{MP} is defined by the expression

$$q \leq m_P(y, x) < q - \Delta q \Rightarrow m_S(y, x) \leftarrow 1, y = \overline{0, Y-1}, x = \overline{0, X-1}.$$

The function F_{RG} is defined by the expressions

$$y \leftarrow y_C(n_S, p_C(n_S)), x \leftarrow x_C(n_S, p_C(n_S)), p_C(n_S) \leftarrow p_C(n_S) - 1$$

$$\begin{aligned} & \forall j, i, k, l (m_S(y+j, x+i)=1) \wedge (m_2(k, l, y, x, j, i) > g) \Rightarrow \\ & \Rightarrow \{m_S(y+j, x+i) \leftarrow m_S(y, x), y_C(n_S, p_C(n_S)) \leftarrow y+j, x_C(n_S, p_C(n_S)) \leftarrow x+i, \\ & m_D(y+j, x+i) \leftarrow m_1(y, x, j, i), p_C(n_S) \leftarrow p_C(n_S)+1, \\ & j \in \{-1, 0, 1\}, i \in \{-1, 0, 1\}, |j|+|i| \neq 0\}, n_S = \overline{2, N_S}, \end{aligned}$$

where y_C and x_C are coordinates of adjacent pixels in matrix M_P from the stack sets $\left\{ Y_C(n_S) = \left\| y_C(n_S, p_C(n_S)) \right\|_{(p_C(n_S)=0, p_C(n_S)-1)} \right\}_{(n_S=\overline{0, N_S})}$ and $\left\{ X_C(n_S) = \left\| x_C(n_S, p_C(n_S)) \right\|_{(p_C(n_S)=0, p_C(n_S)-1)} \right\}_{(n_S=\overline{0, N_S})}$; $m_2(k, l, y, x, j, i) = m_1(k, l, j, i) - m_1(y, x, j, i)$ is the second derivative with respect to intensity calculated on the basis of the first derivatives with respect to intensity for pixels $m_P(y+j, x+i)$, $m_P(y, x)$ and the base pixel $m_P(k, l)$; $m_1(y, x, j, i) = m_P(y, x) - m_P(y+j, x+i)$ is the first derivative with respect to intensity for the boundary pixel $m_P(y, x)$ and the candidate pixel $m_P(y+j, x+i)$; $m_1(k, l, y, x)$ is the first derivative with respect to intensity for the base pixel $m_P(k, l)$ and the boundary pixel $m_P(y, x)$. For all boundary $m_P(y, x)$ and attached $m_P(y+j, x+i)$ pixels $m_1(y, x, j, i) \geq \Delta q$ is performed.

The function F_{SP} is defined by the expression

$$\begin{aligned} m_S(y, x) = 1 \Rightarrow \{N_S \leftarrow N_S + 1, m_S(y, x) \leftarrow N_S, y_S(0) \leftarrow y, x_S(0) \leftarrow x, p_S \leftarrow 1, \\ y_C(N_S, 0) \leftarrow y, x_C(N_S, 0) \leftarrow x, p_C(N_S) \leftarrow 1, F_{PRG}\}, y = \overline{0, Y-1}, x = \overline{0, X-1}, \end{aligned}$$

where $y_S(0)$ and $x_S(0)$ are coordinates of the initial point of growth of the local maximum from the stacks of coordinates $Y_S = \left\| y_S(p_S) \right\|_{(p_S=\overline{0, P_S-1})}$ and $X_S = \left\| x_S(p_S) \right\|_{(p_S=\overline{0, P_S-1})}$; p_S is the pointer to stacks Y_S and X_S .

The function F_{PRG} is executed, while $p_S > 0$ and is defined by the expressions

$$\begin{aligned} & p_S \leftarrow p_S - 1, y \leftarrow y_S(p_S), x \leftarrow x_S(p_S) \forall j, i (m_1(y, x, j, i) = 0) \wedge (m_P(y, x) = q) \Rightarrow \\ & \Rightarrow \{m_S(y+j, x+i) \leftarrow N_S, y_S(p_S) \leftarrow y+j, x_S(p_S) \leftarrow x+i, p_S \leftarrow p_S + 1, \\ & y_C(N_S, p_C(N_S)) \leftarrow y+j, x_C(N_S, p_C(N_S)) \leftarrow x+i, \\ & p_C(N_S) \leftarrow p_C(N_S) + 1, j \in \{-1, 0, 1\}, i \in \{-1, 0, 1\}, |j|+|i| \neq 0\}. \end{aligned}$$

In contrast to the conventional Region Growing (RG) algorithm [11], which uses sequential segment processing, the proposed algorithm expands the boundaries of all regions in a wave-like manner by adding unprocessed adjacent significant pixels, whose brightness satisfies a threshold that is lowered after all significant pixels have been processed. In contrast to the conventional wave Seed Region Growing (SRG) algorithm and its modifications [12–14], which determine all growth points during initialisation, the proposed algorithm selects the initial growth points during the segmentation process and is associated with a threshold value that is gradually reduced from maximum to minimum.

The GWRG algorithm provides incomplete image segmentation into N_S non-overlapping regions, corresponding to objects and the background. The number of regions is unknown in advance and is refined during the segmentation process by gradually lowering the brightness threshold q ($q = 2^B - 1, 0$) with each cycle and detecting new local maxima, which can be single-pixel or multi-pixel (areal). New local maxima and subsequent pixels added to existing regions are detected by comparing with the threshold q . Wave-like growth of regions and weighted separation of regions with smooth boundaries are ensured by monitoring the values of the first and second derivatives of intensity.

The GWRG algorithm can be effectively used as part of the MinLGS algorithm for segmentation based on local minima of the brightness gradient to identify homogeneous regions with high localisation stability under changing weather and time conditions. In this case the initial pixel matrix M_P for the GWRG algorithm should be an inverse gradient image $m_P(y, x) = 255 - i_G(y, x)$, calculating the pixel values of the matrix M_P using the expression $m_P(y, x) = 255 - i_G(y, x)$ for $y = \overline{0, Y-1}, x = \overline{0, X-1}$.

Evaluation of the stability of geometric objects to changes in weather and time conditions for image formation

The examples of test halftone images I_1 and I_2 generated from the same angle but at different times are showed in fig. 1. The images differ in brightness, contrast, shadow distribution and the positions of dynamic objects. Differences in image generation conditions lead to changes in the number and positions of key objects identified by the detectors. As a result, many objects in one image are missing at corresponding coordinates in the other image. As an example, fig. 2 shows the lines identified in images I_1 and I_2 by the Sobel detector [11] and the developed MinLGS algorithm (based on GWRG).



Fig. 1. Test halftone images I_1 and I_2



Fig. 2. Lines highlighted in images I_1 and I_2 by the Sobel detector

To assess the stability of geometric objects to changes in weather and time conditions for image formation dependencies were obtained for the values of the weighted number $\hat{N}_E$ of non-conforming objects calculated by reducing the number $N_E = N_E(1, 2) + N_E(2, 1)$ of non-conforming objects in one $N_E(1, 2)$ and another $N_E(2, 1)$ images to the total number $N_O = N_O(1) + N_O(2)$ of objects in one $N_O(1)$ and another $N_O(2)$ images using the expression $\hat{N}_E = \frac{N_E}{N_O}$.

The weighted characteristics of the detection results are presented in table 1. The table shows that segments extracted by the MinLGS algorithm (based on GWRG) are significantly more resistant to changes in weather and time conditions of image formation compared to points.

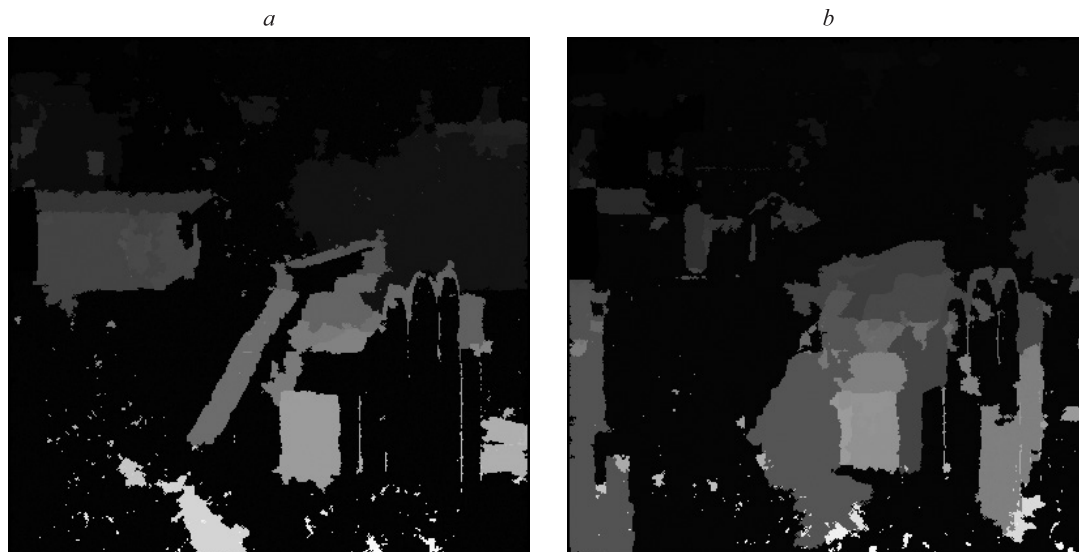


Fig. 3. Segments extracted from images I_1 and I_2 using the MinLGS algorithm (based on GWRG)

**Weighted characteristics of the results
of detection of points and lines of images I_1 and I_2**

Detectors	N_O	$\hat{N}_E$	Source
SIFT	1861	0.960	[15]
SURF	626	0.965	[16]
Sobel	1251	0.951	[11]
Canny	1005	0.961	
MinLGS (based on GWRG)	849	0.124	This work

Conclusions

Algorithms for segmenting grayscale images based on local brightness gradient minima and gradient-wave region growing of local maxima have been developed. It has been established that the developed segmentation algorithms provide more stable region localisation under changing weather and time conditions for image formation, compared to existing line and keypoint extraction algorithms, albeit with higher computational complexity.

References

1. Ince OF, Kim J-S. TIMA SLAM: tracking independently and mapping alto-gether for an uncalibrated multi-camera system. *Sensors*. 2021;21(2):409. DOI: 10.3390/s21020409.2.
2. Hu Z, Fang H, Zhong R, Wei S, Xu B, Dou L. GMP-SLAM: a real-time RGB-D SLAM in dynamic environments using GPU dynamic points detection method. *IFAC-PapersOnLine*. 2023;56(2):5033–5040. DOI: 10.1016/j.ifacol.2023.10.1282.
3. Li D, Shi X, Long Q, Liu S, Yang W, Wang F. DXSLAM: a robust and efficient visual SLAM system with deep features. In: Institute of Electrical and Electronics Engineers. *2020 IEEE/RSJ International conference on intelligent robots and systems (IROS); 2020 October 24 – 2021 January 24; Las Vegas, USA*. [S. l.]: Institute of Electrical and Electronics Engineers; 2020. p. 4958–4965. DOI: 10.1109/IROS45743.2020.9340907.
4. Matsuki H, Scona R, Czarnowski J, Davison AJ. CodeMapping: real-time dense mapping for sparse SLAM using compact scene representations. *IEEE Robotics and Automation Letters*. 2021;6(4):7105–7112. DOI: 10.1109/LRA.2021.3097258.
5. Dai W, Zhang Y, Li P, Fang Z, Scherer S. RGB-D SLAM in dynamic environments using point correlations. *IEEE Transactions on Pattern Analysis & Machine Intelligence*. 2022;44(1):373–389. DOI: 10.1109/TPAMI.2020.3010942.
6. Favorskaya MN. Deep learning for visual SLAM: the state-of-the-art and future trends. *Electronics*. 2023;12(9):1–29. DOI: 10.3390/electronics12092006.
7. Wang S, Hu Q, Zhang X, Li W, Wang Y, Zheng E. LVID-SLAM: a lightweight visual-inertial SLAM for dynamic scenes based on semantic information. *Sensors*. 2025;25(13):4117. DOI: 10.3390/s25134117.

8. Sun Y, Wang Q, Yan C, Feng Y, Tan R, Shi X, et al. D-VINS: dynamic adaptive visual-inertial SLAM with IMU prior and semantic constraints in dynamic scenes. *Remote Sensing*. 2023;15(15):3881. DOI: 10.3390/rs15153881.
9. Zhu D, Liu P, Qiu Q, Wei J, Gong R. BY-SLAM: dynamic visual SLAM system based on BEBLID and semantic information extraction. *Sensors*. 2024;24(14):4693. DOI: 10.3390/s24144693.
10. Al-Furaiji OJM, Nguyen AT, Tsviatkou VYu. A new approach of finding local extrema in grayscale images based on space-oriented masks. *Engineering Letters*. 2020;28(2):586–593.
11. Gonzalez RC, Woods RE. *Digital image processing*. 3rd edition. Delhi: Pearson; 2008. 976 p.
12. Adams R, Bischof L. Seeded region growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1994;16(6):641–647. DOI: 10.1109/34.295913.
13. Fan M, Lee TCM. Variants of seeded region growing. *IET Image Processing*. 2015;9(6):478–485. DOI: 10.1049/iet-ipr.2014.0490.
14. Fan J, Zeng G, Body M, Hacid MS. Seeded region growing: an extensive and comparative study. *Pattern Recognition Letters*. 2005;26(8):1139–1156.
15. Lowe DG. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*. 2004;60(2):91–110.
16. Bay H, Tuytelaars T, van Gool L. SURF: speeded up robust features. In: Leonardis A, Bischof H, Pinz A, editors. *Proceedings of the 9th European conference on computer vision; 2006 May 7–13; Graz, Austria. Part 1*. Berlin: Springer Verlag; 2006. p. 404–417 (Lecture notes in computer science; volume 3951).

Received 24.12.2025 / revised 19.01.2026 / accepted 29.01.2026.

МАТЕМАТИЧЕСКАЯ ЛОГИКА, АЛГЕБРА И ТЕОРИЯ ЧИСЕЛ

MATHEMATICAL LOGIC, ALGEBRA AND NUMBER THEORY

УДК 512.542

О ПРОБЛЕМЕ ШЕМЕТКОВА ОПИСАНИЯ КРИТИЧЕСКИХ ФОРМАЦИЙ В КЛАССЕ τ -ЗАМКНУТЫХ σ -ЛОКАЛЬНЫХ ФОРМАЦИЙ

В. В. СКРУНДЬ¹⁾, И. Н. САФОНОВА¹⁾

¹⁾Белорусский государственный университет, пр. Независимости, 4, 220030, г. Минск, Беларусь

Аннотация. Изучаются минимальные τ -замкнутые σ -локальные формации конечных групп, где σ – некоторое разбиение множества всех простых чисел $\mathbb{P}$, τ – произвольный подгрупповой функтор. Получено описание минимальных τ -замкнутых σ -локальных не $\mathfrak{H}$ -формаций для произвольной σ -локальной формации классического типа, т. е. такой σ -локальной формации, которая имеет σ -локальное определение, все неабелевы значения которого являются σ -локальными формациями. Тем самым решена задача об описании критических формаций в классе τ -замкнутых σ -локальных формаций, предложенная Л. А. Шеметковым (1980).

Ключевые слова: конечная группа; подгрупповой функтор; τ -замкнутая формация; формационная σ -функция; σ -локальная формация; критическая σ -локальная формация; формация классического типа.

Благодарность. Исследование выполнено при поддержке Министерства образования Республики Беларусь (грант № 20211328). Авторы выражают признательность рецензенту за внимательное прочтение статьи.

Образец цитирования:

Скрудзь ВВ, Сафонова ИН. О проблеме Шеметкова описания критических формаций в классе τ -замкнутых σ -локальных формаций. *Журнал Белорусского государственного университета. Математика. Информатика.* 2026;1:163–178 (на англ.).
EDN: WDVFOF

For citation:

Skrundz VV, Safonova IN. On Shemetkov's problem of describing critical formations in the class of τ -closed σ -local formations. *Journal of the Belarusian State University. Mathematics and Informatics.* 2026;1:163–178.
EDN: WDVFOF

Авторы:

Валентина Викторовна Скрудзь – аспирантка кафедры высшей алгебры и защиты информации механико-математического факультета. Научный руководитель – И. Н. Сафонова.

Инна Николаевна Сафонова – доктор физико-математических наук, доцент; профессор кафедры высшей алгебры и защиты информации механико-математического факультета.

Authors:

Valentina V. Skrundz, postgraduate student at the department of higher algebra and information security, faculty of mechanics and mathematics.
vallik@mail.ru

Inna N. Safonova, doctor of science (physics and mathematics), docent; professor at the department of higher algebra and information security, faculty of mechanics and mathematics.
safonova@bsu.by

<https://orcid.org/0000-0001-6896-7208>



ON SHEMETKOV'S PROBLEM OF DESCRIBING CRITICAL FORMATIONS IN THE CLASS OF τ -CLOSED σ -LOCAL FORMATIONS

V. V. SKRUNDZ^a, I. N. SAFONOVA^a

^aBelarusian State University, 4 Niezaliezhnasci Avenue, Minsk 220030, Belarus

Corresponding author: I. N. Safonova (safonova@bsu.by)

Abstract. The critical τ -closed σ -local formations of finite groups are studied, where σ is some partition of the set of all primes $\mathbb{P}$, τ is an arbitrary subgroup functor. A description of minimal τ -closed σ -local non- $\mathfrak{H}$ -formations is obtained for an arbitrary σ -local formation of classical type, i. e. a σ -local formation that has a σ -local definition all of whose non-Abelian values are σ -local formations. Thus, the problem of describing critical formations in the class of τ -closed σ -local formations proposed by L. A. Shemetkov (1980) is solved.

Keywords: finite group; subgroup functor; τ -closed formation; formation σ -function; σ -local formation; critical σ -local formation; formation of classical type.

Acknowledgements. This research was supported by the Ministry of Education of the Republic of Belarus (grant No. 20211328). The authors thank the referee for careful reading of the paper.

Introduction

The problem of studying critical formations was posed by L. A. Shemetkov [1] at the 6th All-Union symposium on group theory. In the theory of local formations, a minimal local non- $\mathfrak{H}$ -formation [1] or $\mathfrak{H}$ -critical formation [2] is defined as a local formation $\mathfrak{F} \not\leq \mathfrak{H}$, all proper local subformations of which are contained in the class of groups $\mathfrak{H}$. The solution to L. A. Shemetkov's problem for local formations was obtained by A. N. Skiba in a series of works from 1980–1993 and culminated in the construction of the theory of critical local formations. The most general result was the description of minimal local non- $\mathfrak{H}$ -formations for the case when $\mathfrak{H}$ is an arbitrary formation of classical type [3], i. e. a formation that has a local screen all of whose non-Abelian values are local. The theory of τ -closed σ -local formations was constructed by A. N. Skiba in the monograph [4]. Results on critical formations were widely used in studying the internal structure and classification of local and τ -closed local formations, as well as in studying their irreducible factorisations [4; 5].

The development of the theory of σ -local formations has led to the necessity of studying and classifying critical σ -local and functor-closed or, in other words, τ -closed σ -local formations, where τ is a subgroup functor in the sense of A. N. Skiba [4]. In this context, following [1; 2], a *minimal τ -closed σ -local non- $\mathfrak{H}$ -formation* or $\mathfrak{H}_\sigma^\tau$ -critical formation is defined as a τ -closed σ -local formation $\mathfrak{F} \not\leq \mathfrak{H}$, all proper τ -closed σ -local subformations of which are contained in the class of groups $\mathfrak{H}$. It should be noted that L. A. Shemetkov's problem of classifying critical formations in the class of σ -local formations was solved in the works [6; 7] for an arbitrary σ -local formation $\mathfrak{H}$ of classical type. Here, a formation $\mathfrak{H}$ is called [7] a *σ -local formation of classical type* if it has a σ -local definition all of whose non-Abelian values are σ -local.

The study of $\mathfrak{H}_\sigma^\tau$ -critical formations was initiated by the authors in the work [8], where a criterion for minimal τ -closed σ -local non- $\mathfrak{H}$ -formations was proved, and a description of such formations was given for a number of σ -local formations of classical type, in particular, for formations of all σ -nilpotent and all σ -soluble groups.

The main result of this paper is the following theorem, which provides a solution to the aforementioned problem of L. A. Shemetkov in the class of τ -closed σ -local formations.

Theorem 1. *Let $\mathfrak{H}$ be a σ -local formation of classical type and h be its canonical σ -local definition. Then $\mathfrak{F}$ is a minimal τ -closed σ -local non- $\mathfrak{H}$ -formation if and only if $\mathfrak{F} = I_\sigma^\tau \text{ form } G$, where G is a monolithic $\bar{\tau}$ -minimal non- $\mathfrak{H}$ -group with monolith $P = G^\mathfrak{H}$, and one of the following conditions holds:*

- 1) $G = P$ is a simple σ_i -group such that $\sigma_i \notin \sigma(\mathfrak{H})$ and $\tau(G) \subseteq \{1, G\}$;
- 2) P is a non- σ -primary group and G is a $\bar{\tau}$ -minimal non- $h(\sigma_i)$ -group with $P = G^{h(\sigma_i)}$ for all $\sigma_i \in \sigma(P)$;
- 3) $G = P \rtimes K$, where $P = C_G(P)$ is a p -group, $p \in \sigma_i$, and K is either a monolithic $\bar{\tau}$ -minimal non- $h(\sigma_i)$ -group with monolith $Q = K^{h(\sigma_i)} \not\leq \Phi(K)$, where $\sigma_i \notin \sigma(Q)$, or a minimal non- $h(\sigma_i)$ -group of one of the following types:
 - a) the quaternion group of order 8, if $2 \notin \sigma_i$;
 - b) an extraspecial group of order q^3 of prime odd exponent $q \notin \sigma_i$;
 - c) a cyclic q -group, $q \notin \sigma_i$.

Notations and definitions

The main definitions, notations of the theory of σ -properties of groups and general properties of σ -local formations are presented in the works [9–16].

Let $\sigma = \{\sigma_i | i \in I\}$ be some partition of the set of all primes $\mathbb{P}$, i. e. $\sigma = \{\sigma_i | i \in I\}$, where $\mathbb{P} = \bigcup_{i \in I} \sigma_i$ and $\sigma_i \cap \sigma_j = \emptyset$ for all $i \neq j$, G be a group, $\mathfrak{F}$ be a class of groups. Then $\sigma(G) = \{\sigma_i | \sigma_i \cap \pi(G) \neq \emptyset\}$; $\sigma(\mathfrak{F}) = \bigcup_{G \in \mathfrak{F}} \sigma(G)$.

A group G is called [9]: σ -primary, if G is a σ_i -group for some i ; σ -nilpotent, if $G = G_1 \times \dots \times G_t$ for some σ -primary groups $G_1, \dots, G_t$; σ -soluble, if $G = 1$ or $G \neq 1$ and every chief factor of G is σ -primary.

By $\mathfrak{S}_\sigma, \mathfrak{N}_\sigma, \mathfrak{S}_{\sigma_i}$ and (1) we denote the classes of all σ -soluble, σ -nilpotent, σ_i -groups and identity groups, respectively.

Let $\mathfrak{F}$ be a non-empty formation. Then for any group G the symbol $G^\mathfrak{F}$ denotes the $\mathfrak{F}$ -residual of the group G , i. e. the smallest normal subgroup of G such that the quotient group belongs to the formation $\mathfrak{F}$. If $\mathfrak{F}$ and $\mathfrak{H}$ are formations, then their product $\mathfrak{F}\mathfrak{H}$ is defined as follows: $\mathfrak{F}\mathfrak{H} = (G | G^\mathfrak{H} \in \mathfrak{F})$.

The symbol $O_{\sigma_i, \sigma_i}(G)$ denotes the $(\mathfrak{S}_{\sigma_i}, \mathfrak{S}_{\sigma_i})$ -radical of the group G (the largest normal $(\mathfrak{S}_{\sigma_i}, \mathfrak{S}_{\sigma_i})$ -subgroup of the group G), i. e. $O_{\sigma_i, \sigma_i}(G) = G_{\mathfrak{S}_{\sigma_i}, \mathfrak{S}_{\sigma_i}}$.

A function f of the form $f: \sigma \rightarrow \{\text{formations of groups}\}$ is called [10] a formation σ -function. For every formation σ -function f the class $LF_\sigma(f)$ is defined as follows:

$$LF_\sigma(f) = (G | G = 1 \text{ or } G \neq 1 \text{ and } G/O_{\sigma_i, \sigma_i}(G) \in f(\sigma_i) \text{ for all } \sigma_i \in \sigma(G)).$$

If for some formation σ -function f we have $\mathfrak{F} = LF_\sigma(f)$, then it is said that the formation $\mathfrak{F}$ is σ -local, and f is a σ -local definition of $\mathfrak{F}$.

Suppose that to each group G a certain system of its subgroups $\tau(G)$ is assigned. Then τ is called a subgroup functor [4, p. 16], if $G \in \tau(G)$ and for any epimorphism $\varphi: A \rightarrow B$ and for any groups $H \in \tau(A)$ and $T \in \tau(B)$ we have $H^\varphi \in \tau(B)$ and $H^{\varphi^{-1}} \in \tau(A)$. A class of groups $\mathfrak{F}$ is called τ -closed, if $\tau(G) \subseteq \mathfrak{F}$ for any group $G \in \mathfrak{F}$.

A subgroup functor τ is called [4, p. 16]: closed, if for any two groups G and $H \in \tau(G)$ we have $\tau(H) \subseteq \tau(G)$; trivial, if for any group G we have $\tau(G) = \{G\}$; identity, if for any group G we have $\tau(G)$ is the set of all subgroups of the group G .

On the set of all subgroup functors, a partial order ($\leq$) is introduced [4, p. 19], assuming that $\tau_1 \leq \tau_2$ holds if and only if for any group G the inclusion $\tau_1(G) \subseteq \tau_2(G)$ is valid. For any collection of subgroup functors $\{\tau_i | i \in I\}$ their intersection $\tau = \bigcap_{i \in I} \tau_i$ is defined: $\tau(G) = \bigcap_{i \in I} \tau_i(G)$ for any group G .

Let now τ be an arbitrary subgroup functor, $\bar{\tau}$ be the intersection of all such closed subgroup functors τ_i , for which $\tau \leq \tau_i$. The functor $\bar{\tau}$ is called the closure of the functor τ [4, p. 20].

For any set of groups $\mathfrak{X}$ the symbol $S_\tau \mathfrak{X}$ [4, p. 23] denotes the set of all groups H , such that $H \in \tau(G)$ for some group $G \in \mathfrak{X}$. It is clear that a class of groups $\mathfrak{F}$ is τ -closed, if $S_\tau(\mathfrak{F}) = \mathfrak{F}$.

If f is a formation σ -function, then by $\text{Supp}(f)$ we denote the support of f , i. e. the set of all σ_i such that $f(\sigma_i) \neq \emptyset$.

A formation σ -function f is called: τ -valued, if $f(\sigma_i)$ is a τ -closed formation for all $\sigma_i \in \text{Supp}(f)$; integrated, if $f(\sigma_i) \subseteq LF_\sigma(f)$ for all i ; full, if $f(\sigma_i) = \mathfrak{S}_{\sigma_i} f(\sigma_i)$ for all i . If f is a full integrated formation σ -function and $\mathfrak{F} = LF_\sigma(f)$, then f is called the canonical σ -local definition of $\mathfrak{F}$. The smallest σ -local definition of $\mathfrak{F}$ is its σ -local definition f such that $f(\sigma_i) \subseteq h(\sigma_i)$ for any σ -local definition h of $\mathfrak{F}$ and all i .

Let $\mathfrak{H}$ be an arbitrary class of groups. A group G is called [4, p. 38] a τ -minimal non- $\mathfrak{H}$ -group or a $\mathfrak{H}^\tau$ -critical group, if $G \notin \mathfrak{H}$, but every proper τ -subgroup of the group G belongs to the class $\mathfrak{H}$.

For an arbitrary set of groups $\mathfrak{X}$ and any $\sigma_i \in \sigma$ the symbol $\mathfrak{X}(\sigma_i)$ [13, p. 962] denotes the class of groups defined as follows: $\mathfrak{X}(\sigma_i) = (G/O_{\sigma_i, \sigma_i}(G) | G \in \mathfrak{X})$, if $\sigma_i \in \sigma(\mathfrak{X})$, $\mathfrak{X}(\sigma_i) = \emptyset$, if $\sigma_i \notin \sigma(\mathfrak{X})$.

The set of all τ -closed σ -local formations is denoted by L_σ^τ . Formations from L_σ^τ are called L_σ^τ -formations. For any set of groups $\mathfrak{M}$ the symbol $L_\sigma^\tau \text{ form } \mathfrak{M}$ denotes the τ -closed σ -local formation generated by the class of groups $\mathfrak{M}$, i. e. the intersection of all L_σ^τ -formations containing $\mathfrak{M}$. In the case when $\mathfrak{M} = \{G\}$ the formation $L_\sigma^\tau \text{ form } G$ is called a one-generated L_σ^τ -formation.

Let now $\{\mathfrak{F}_j \mid j \in J\}$ be some set of τ -closed σ -local formations. Then we set $\vee_{\sigma}^{\tau}(\mathfrak{F}_j \mid j \in J) = l_{\sigma}^{\tau} \text{form}(\cup_{j \in J} \mathfrak{F}_j)$. In particular, if $\mathfrak{H}$ and $\mathfrak{F}$ are τ -closed σ -local formations, then $\mathfrak{H} \vee_{\sigma}^{\tau} \mathfrak{F} = l_{\sigma}^{\tau} \text{form}(\mathfrak{H} \cup \mathfrak{F})$.

Let $\mathfrak{F}$ be a τ -closed σ -local formation. The formation $\mathfrak{F}$ is called an *irreducible τ -closed σ -local formation* or an *l_{σ}^{τ} -irreducible formation*, if $\mathfrak{F} \neq l_{\sigma}^{\tau} \text{form}(\cup_{i \in I} \mathfrak{X}_i) = \vee_{\sigma}^{\tau}(\mathfrak{X}_i \mid i \in I)$, where $\{\mathfrak{X}_i \mid i \in I\}$ is the set of all proper τ -closed σ -local subformations of $\mathfrak{F}$. If there exist proper τ -closed σ -local subformations $\mathfrak{X}$ and $\mathfrak{H}$ of $\mathfrak{F}$, such that $\mathfrak{F} = \mathfrak{X} \vee_{\sigma}^{\tau} \mathfrak{H}$, then the formation $\mathfrak{F}$ is called a *reducible τ -closed σ -local formation* or *l_{σ}^{τ} -reducible formation*.

Auxiliary results

For the proof of the main result of the paper, we will need the following known facts from formation theory.

Lemma 1 [13, lemma 2.1]. *Let f and h be formation σ -functions and let $\Pi = \text{Supp}(f)$. Suppose that $\mathfrak{F} = LF_{\sigma}(f) = LF_{\sigma}(h)$. Then:*

- (1) $\Pi = \sigma(\mathfrak{F})$;
- (2) $\mathfrak{F} = \left(\bigcap_{\sigma_i \in \Pi} \mathfrak{G}_{\sigma_i} \mathfrak{G}_{\sigma_i} f(\sigma_i) \right) \cap \mathfrak{G}_{\Pi}$. Consequently, $\mathfrak{F}$ is a saturated formation;
- (3) if every group in $\mathfrak{F}$ is σ -soluble, then $\mathfrak{F} = \left(\bigcap_{\sigma_i \in \Pi} \mathfrak{G}_{\sigma_i} \mathfrak{G}_{\sigma_i} f(\sigma_i) \right) \cap \mathfrak{G}_{\Pi}$;
- (4) if $\sigma_i \in \Pi$, then $\mathfrak{G}_{\sigma_i}(f(\sigma_i) \cap \mathfrak{F}) = \mathfrak{G}_{\sigma_i}(h(\sigma_i) \cap \mathfrak{F}) \subseteq \mathfrak{F}$;
- (5) $\mathfrak{F} = LF_{\sigma}(F)$, where F is the unique formation σ -function such that $F(\sigma_i) = \mathfrak{G}_{\sigma_i} F(\sigma_i) \subseteq \mathfrak{F}$ for all $\sigma_i \in \Pi$ and $F(\sigma_i) = \emptyset$ for all $\sigma_i \in \Pi'$. Moreover, $F(\sigma_i) = \mathfrak{G}_{\sigma_i}(f(\sigma_i) \cap \mathfrak{F})$ for all i .

Lemma 2 [17, lemma 1.1 (a)]. *Let $\mathfrak{F}$ be a Fitting class and G be a finite group. If $N \triangleleft\triangleleft G$, then $N_{\mathfrak{F}} = N \cap G_{\mathfrak{F}}$.*

Lemma 3 [18, lemma 1.2]. *Let $\mathfrak{F}$ be a non-empty formation, $K \trianglelefteq G$. Then the following statements hold:*

- (1) $(G/K)^{\mathfrak{F}} = G^{\mathfrak{F}}K/K$;
- (2) if $G = HK$, then $H^{\mathfrak{F}}K = G^{\mathfrak{F}}K$;
- (3) if $G = HK$ and $K \subseteq G^{\mathfrak{F}}$, then $H^{\mathfrak{F}}K = G^{\mathfrak{F}}$.

Lemma 4 [16, lemma 22]. *Let $\mathfrak{M} = LF_{\sigma}(m)$, where m is a formation σ -function such that $m(\sigma_i)$ is a τ -closed formation for all $\sigma_i \in \text{Supp}(m)$. Then the formation $\mathfrak{M}$ is τ -closed.*

Recall that $\tau \text{form } \mathfrak{X}$ denotes the τ -closed formation generated by the class of groups $\mathfrak{X}$, i. e. the intersection of all τ -closed formations containing $\mathfrak{X}$.

Lemma 5 [15, theorem 1.1]. *Let $\mathfrak{X}$ be some non-empty set of groups, $\mathfrak{F} = l_{\sigma}^{\tau} \text{form } \mathfrak{X} = LF_{\sigma}(f)$, where f is the smallest τ -valued σ -local definition of $\mathfrak{F}$. Then the following statements hold:*

- (1) $\sigma(\mathfrak{X}) = \sigma(\mathfrak{F})$;
- (2) $f(\sigma_i) = \tau \text{form}(\mathfrak{X}(\sigma_i)) = \tau \text{form}(\mathfrak{F}(\sigma_i))$ for all i ;
- (3) if h is an arbitrary τ -valued σ -local definition of $\mathfrak{F}$, then for all $\sigma_i \in \sigma(\mathfrak{X})$ we have $f(\sigma_i) = \tau \text{form}(A \mid A \in \mathfrak{F} \cap h(\sigma_i), O_{\sigma_i}(A) = 1)$.

Following [18], we define a partial order ($\leq$) on the set of all formation σ -functions as follows: if f_1 and f_2 are formation σ -functions, then $f_1 \leq f_2$, if $f_1(\sigma_i) \subseteq f_2(\sigma_i)$ for any $\sigma_i \in \sigma$.

The following lemma is a special case of corollary 3.1 from paper [15].

Lemma 6 [15, corollary 3.1]. *Let $\mathfrak{F}_j = LF_{\sigma}(f_j)$, where f_j is the smallest τ -valued σ -local definition of the formation $\mathfrak{F}_j$, $j = 1, 2$. Then $\mathfrak{F}_1 \subseteq \mathfrak{F}_2$ if and only if $f_1 \leq f_2$.*

Lemma 7 [14, lemma 11]. *If $\mathfrak{F} = LF_{\sigma}(f)$ and $G/O_{\sigma_i}(G) \in f(\sigma_i) \cap \mathfrak{F}$ for some $\sigma_i \in \sigma(G)$, then $G \in \mathfrak{F}$.*

A τ -closed formation $\mathfrak{F}$ is called [4] an *irreducible τ -closed formation* or a *τ -irreducible formation*, if $\mathfrak{F} \neq \tau \text{form}(\cup_{i \in I} \mathfrak{X}_i)$, where $\{\mathfrak{X}_i \mid i \in I\}$ is the set of all proper τ -closed subformations of $\mathfrak{F}$.

Lemma 8 [4, lemma 2.1.5]. *Let A be a monolithic group with monolith $R \not\subseteq \Phi(A)$. Then the formation $\mathfrak{F} = \tau \text{form } A$ is τ -irreducible and it has a unique maximal τ -closed subformation $\mathfrak{M} = \tau \text{form}(\{A/R\} \cup \mathfrak{X})$, where $\mathfrak{X}$ is the set of all proper τ -subgroups of A .*

Lemma 9 [7, lemma 2.1]. *Let $\emptyset \neq \Pi \subseteq \sigma$. Then the class $\mathfrak{G}_{\Pi}$ of all Π -groups and the class $\mathfrak{N}_{\Pi}$ of all σ -nilpotent Π -groups are σ -local formations and the following statements hold:*

- (1) $\mathfrak{G}_{\Pi} = LF_{\sigma}(g)$, where g is the canonical σ -local definition of the formation $\mathfrak{G}_{\Pi}$. Here, $g(\sigma_i) = \mathfrak{G}_{\Pi}$ for all $\sigma_i \in \Pi$ and $g(\sigma_i) = \emptyset$ for all $\sigma_i \in \Pi'$;

(2) $\mathfrak{N}_\Pi = LF_\sigma(n) = LF_\sigma(k)$, where n and k , respectively, are the smallest and canonical σ -local definitions of the formation $\mathfrak{N}_\Pi$. Here, $n(\sigma_i) = (1)$ for all $\sigma_i \in \Pi$ and $n(\sigma_i) = \emptyset$ for all $\sigma_i \in \Pi'$, $k(\sigma_i) = \mathfrak{G}_{\sigma_i}$ for all $\sigma_i \in \Pi$ and $k(\sigma_i) = \emptyset$ for all $\sigma_i \in \Pi'$.

Lemma 10 [13, theorem 1.14]. Let $\mathfrak{M} = LF_\sigma(m)$, where m is integrated, $\Pi = \sigma(\mathfrak{M})$, and let $\mathfrak{H}$ be a non-empty formation. Suppose that either (i) $\mathfrak{H} = LF_\sigma(h)$, where h is integrated, or (ii) $\mathfrak{G}_{\sigma_i} \mathfrak{H} \subseteq \mathfrak{H}$ for all $\sigma_i \in \sigma \setminus \Pi$. Then $\mathfrak{M}\mathfrak{H} = LF_\sigma(f)$, where

$$f(\sigma_i) = \begin{cases} m(\sigma_i)\mathfrak{H}, & \text{if } \sigma_i \in \Pi, \\ h(\sigma_i), & \text{if } \sigma_i \in \sigma \setminus \Pi \text{ in case (i),} \\ \mathfrak{H}, & \text{if } \sigma_i \in \sigma \setminus \Pi \text{ in case (ii).} \end{cases}$$

Lemma 11 [8, theorem 3.1]. Let $\mathfrak{H} = LF_\sigma(h)$, $\mathfrak{F} = LF_\sigma(f)$, where h is the canonical σ -local definition of $\mathfrak{H}$, f is the smallest τ -valued σ -local definition of $\mathfrak{F}$. Then $\mathfrak{F}$ is a minimal τ -closed σ -local non- $\mathfrak{H}$ -formation if and only if $\mathfrak{F} = l_\sigma^\tau \text{form } G$, where G is a monolithic τ -minimal non- $\mathfrak{H}$ -group with monolith $P = G^\mathfrak{H}$, such that for all $\sigma_i \in \sigma(P)$ the formation $f(\sigma_i)$ is a minimal τ -closed non- $h(\sigma_i)$ -formation.

Lemma 12 [5, corollary 7.14]. The set of n -multiply local formations is a subsemigroup in the semigroup $G\mathfrak{S}$ of all formations.

Lemma 13 [5, lemma 18.8]. If a group G has only one minimal normal subgroup and $O_p(G) = 1$ (p is some prime), then there exists a faithful irreducible $F_p G$ -module, where F_p is the field of p elements.

Lemma 14 [4, corollary 1.2.23]. Let $\mathfrak{X}$ be an arbitrary set of groups, $\mathfrak{M} = S_\tau(\mathfrak{X})$. Then $\tau \text{form } \mathfrak{X} = \text{form } \mathfrak{M}$. The following lemma is a consequence of lemmas 1 and 6.

Lemma 15. Let $\mathfrak{F}_i = LF_\sigma(F_i)$, where F_i is the canonical σ -local definition of $\mathfrak{F}_i$, $i = 1, 2$. Then $\mathfrak{F}_1 \subseteq \mathfrak{F}_2$ if and only if $F_1 \leq F_2$.

Lemma 16 [5, corollary 19.14]. A group is nilpotent if and only if all subformations of the formation it generates are hereditary.

Lemma 17 [4, lemma 2.2.4]. A formation $\mathfrak{F}$ is a minimal τ -closed non-Abelian formation if and only if $\mathfrak{F} = \tau \text{form } A$, where A is a τ -minimal non-Abelian group and one of the following conditions holds:

- 1) A is a monolithic group with a non-Frattini monolith $R = A'$;
- 2) A is the quaternion group of order 8;
- 3) A is a nilpotent group of class 2 and its exponent is 4;
- 4) A is an extraspecial p -group of order p^3 and its exponent is $p > 2$;
- 5) A is a nilpotent group of class 2 having a prime odd exponent.

Recall that a group $G \in R_0\mathfrak{X}$ [17, p. 264] if and only if G has normal subgroups $N_1, \dots, N_t$ ($t \geq 2$) such that $G/N_i \in \mathfrak{X}$ and $\bigcap_{i=1}^t N_i = 1$. The set of all homomorphic images of all groups from $\mathfrak{X}$ is denoted by $H\mathfrak{X}$.

Lemma 18 [4, lemma 1.2.22]. For any set of groups $\mathfrak{X}$ the following equality holds:

$$\tau \text{form } \mathfrak{X} = HR_0 S_\tau(\mathfrak{X}).$$

Lemma 19 [5, lemma 8.12]. Let $A \in s \text{form } G$, where G is a finite group. Then the following statements hold:

- (1) the exponent of the group A does not exceed the exponent of the group G ;
- (2) every chief factor of the group A is isomorphic to some chief factor of the group G ;
- (3) every composition factor of the group A is isomorphic to some composition factor of the group G ;
- (4) the class of any nilpotent factor of the group A does not exceed the maximum of the classes of the nilpotent factors of the group G .

Lemma 20 [12, lemma 2.4]. If $\mathfrak{F} = LF_\sigma(f)$, then $\mathfrak{F} = LF_\sigma(t)$, where $t(\sigma_i) = f(\sigma_i) \cap \mathfrak{F}$ for all $\sigma_i \in \sigma$.

Lemma 21 [8, theorem 4.1]. Let $\mathfrak{F}$ be a τ -closed σ -local formation, $\mathfrak{F} \not\subseteq \mathfrak{G}_\Pi$, where $\emptyset \neq \Pi \subseteq \sigma$. Then $\mathfrak{F}$ is a minimal τ -closed σ -local non- $\mathfrak{G}_\Pi$ -formation if and only if $\mathfrak{F} = \mathfrak{G}_{\sigma_i} l_\sigma^\tau \text{form } G$, where G is a simple σ_i -group such that $\sigma_i \notin \Pi$ and $\tau(G) \subseteq \{1, G\}$.

Main result

Before we prove theorem 1, we establish the validity of the following lemmas.

Lemma 22. Let $\mathfrak{F} = LF_\sigma(f)$, where f is the canonical σ -local definition of $\mathfrak{F}$ and let $\mathfrak{H}$ be a non-empty formation such that $\sigma(\mathfrak{H}) \subseteq \sigma(\mathfrak{F})$. Then $\mathfrak{F}\mathfrak{H} = LF_\sigma(h)$, where h is the canonical σ -local definition of $\mathfrak{F}\mathfrak{H}$, and the following statements hold:

- (1) $h(\sigma_i) = f(\sigma_i)\mathfrak{H}$, if $\sigma_i \in \sigma(\mathfrak{F})$;
- (2) $h(\sigma_i) = \emptyset$, if $\sigma_i \notin \sigma(\mathfrak{F})$.

Proof. First, note that $h(\sigma_i) \subseteq \mathfrak{F}\mathfrak{H}$ for any $\sigma_i \in \sigma(\mathfrak{F})$. Let $\mathfrak{M} = LF_\sigma(h)$. By lemma 1 (1) we have $\sigma(\mathfrak{M}) = \sigma(\mathfrak{F})$.

Assume that $\mathfrak{M} \not\subseteq \mathfrak{F}\mathfrak{H}$ and let G be a group of minimal order from $\mathfrak{M} \setminus \mathfrak{F}\mathfrak{H}$. Then G is a monolithic group with monolith $P = G^{\mathfrak{F}\mathfrak{H}}$.

First, consider the case when P is a σ -primary group. Then P is a σ_i -group for some $\sigma_i \in \sigma(\mathfrak{F})$ and $O_{\sigma_i, \sigma_i}(G) = O_{\sigma_i}(G)$. Since $G \in \mathfrak{M}$, we have

$$G/O_{\sigma_i}(G) = G/O_{\sigma_i, \sigma_i}(G) \in h(\sigma_i) = f(\sigma_i)\mathfrak{H} = \mathfrak{G}_{\sigma_i}f(\sigma_i)\mathfrak{H}.$$

Therefore, $G \in f(\sigma_i)\mathfrak{H} \subseteq \mathfrak{F}\mathfrak{H}$, a contradiction. Hence, the group P is non- σ -primary. Then $O_{\sigma_i, \sigma_i}(G) = 1$ for any $\sigma_i \in \sigma(P)$. Consequently, $G \simeq G/O_{\sigma_i, \sigma_i}(G) \in f(\sigma_i) \subseteq \mathfrak{F}\mathfrak{H}$ and we again obtain a contradiction with the choice of the group G . Hence, this case is impossible. Therefore, $\mathfrak{M} \subseteq \mathfrak{F}\mathfrak{H}$.

Now suppose that $\mathfrak{F}\mathfrak{H} \not\subseteq \mathfrak{M}$ and let A be a group of minimal order from $\mathfrak{F}\mathfrak{H} \setminus \mathfrak{M}$. Then A is a monolithic group with monolith $R = A^{\mathfrak{M}}$. Since $A \in \mathfrak{F}\mathfrak{H}$, we have $A^{\mathfrak{H}} \in \mathfrak{F}$. Also note that since $\mathfrak{H} \subseteq \mathfrak{F}\mathfrak{H}$, we have $A^{\mathfrak{H}} \neq 1$. Therefore, $R \leq A^{\mathfrak{H}}$. Since $A^{\mathfrak{H}} \in \mathfrak{F}$, we have $A^{\mathfrak{H}}/O_{\sigma_i, \sigma_i}(A^{\mathfrak{H}}) \in f(\sigma_i)$. Since $O_{\sigma_i, \sigma_i}(A^{\mathfrak{H}}) = O_{\sigma_i, \sigma_i}(A) \cap A^{\mathfrak{H}}$ by lemma 2, we have

$$O_{\sigma_i, \sigma_i}(A)A^{\mathfrak{H}}/O_{\sigma_i, \sigma_i}(A) \simeq A^{\mathfrak{H}}/O_{\sigma_i, \sigma_i}(A) \cap A^{\mathfrak{H}} = A^{\mathfrak{H}}/O_{\sigma_i, \sigma_i}(A^{\mathfrak{H}}) \in f(\sigma_i).$$

But by lemma 3 (1) we have the equality $O_{\sigma_i, \sigma_i}(A)A^{\mathfrak{H}}/O_{\sigma_i, \sigma_i}(A) = (A/O_{\sigma_i, \sigma_i}(A))^{\mathfrak{H}}$. Therefore, $A/O_{\sigma_i, \sigma_i}(A) \in f(\sigma_i)\mathfrak{H} = h(\sigma_i)$ for any $\sigma_i \in \sigma(R)$. Moreover, since $A/R \in \mathfrak{M}$ and $O_{\sigma_i, \sigma_i}(A/R) = O_{\sigma_i, \sigma_i}(A)/R$ for any $\sigma_i \in \sigma(A) \setminus \sigma(R)$, we have

$$A/O_{\sigma_i, \sigma_i}(A) \simeq (A/R)/O_{\sigma_i, \sigma_i}(A/R) = (A/R)/O_{\sigma_i, \sigma_i}(A/R) \in h(\sigma_i)$$

for any $\sigma_i \in \sigma(A) \setminus \sigma(R)$. Thus, $A/O_{\sigma_i, \sigma_i}(A) \in h(\sigma_i)$ for any $\sigma_i \in \sigma(A)$. Consequently, $A \in \mathfrak{M}$. The obtained contradiction shows that our assumption is false. Hence, $\mathfrak{F}\mathfrak{H} \subseteq \mathfrak{M}$ and therefore $\mathfrak{F}\mathfrak{H} = LF_\sigma(h)$. Since f is the canonical σ -local definition of the formation $\mathfrak{F}$, for any σ_i the equality

$$\mathfrak{G}_{\sigma_i}h(\sigma_i) = \mathfrak{G}_{\sigma_i}(f(\sigma_i)\mathfrak{H}) = (\mathfrak{G}_{\sigma_i}f(\sigma_i))\mathfrak{H} = f(\sigma_i)\mathfrak{H} = h(\sigma_i)$$

holds. Since also $h(\sigma_i) \subseteq LF_\sigma(h) = \mathfrak{F}\mathfrak{H}$ for any σ_i , it follows that h is the canonical σ -local definition of the formation $\mathfrak{F}\mathfrak{H}$. The lemma is proved.

Lemma 23. Let $G = P \rtimes H$, where $P = C_G(P)$ is a p -group, $p \in \sigma_i$, and H is a monolithic group with monolith Q , such that $\sigma_i \notin \sigma(Q)$ and $\Phi(H) = 1$. And let $\mathfrak{X}$ be the set of all proper τ -subgroups of H . Then $\mathfrak{F} = l_\sigma^\tau \text{ form } G$ is l_σ^τ -irreducible and its maximal τ -closed σ -local subformation $\mathfrak{H}$ has an integrated τ -valued σ -local definition h such that $h(\sigma_i) = \tau \text{ form}(\mathfrak{X} \cup \{H/Q\})$; $h(\sigma_j) = \tau \text{ form}(G/O_{\sigma_j, \sigma_j}(G))$ for all $\sigma_j \in \sigma(G) \setminus \{\sigma_i\}$ and $h(\sigma_j) = \emptyset$ for all $\sigma_j \notin \sigma(G)$.

Proof. Let $\mathfrak{H} = LF_\sigma(h)$, where h is the formation σ -function from the condition of the lemma. Then by lemma 4 the formation $\mathfrak{H}$ is τ -closed. Let us show that $G/P \in \mathfrak{H}$. Let $\sigma_j \in \sigma(G) \setminus \{\sigma_i\}$. Then $P \subseteq O_{\sigma_j}(G)$. Hence, $O_{\sigma_j, \sigma_j}(G/P) = O_{\sigma_j, \sigma_j}(G)/P$. Therefore,

$$G/O_{\sigma_j, \sigma_j}(G) \simeq (G/P)/(O_{\sigma_j, \sigma_j}(G)/P) = (G/P)/O_{\sigma_j, \sigma_j}(G/P).$$

But by the construction of h we have $h(\sigma_j) = \tau \text{ form}(G/O_{\sigma_j, \sigma_j}(G))$. Consequently,

$$(G/P)/O_{\sigma_j, \sigma_j}(G/P) \in h(\sigma_j).$$

Let under the canonical isomorphism $H \simeq G/P$ the subgroup Q in the group G/P correspond to the subgroup Q_1/P . Then, by the condition of the lemma, $\sigma_i \in \sigma(Q_1/P)$. Hence, $Q_1/P \subseteq O_{\sigma_i}(G/P)$. Consequently, $Q_1/P \subseteq O_{\sigma_i}(G/P) \subseteq O_{\sigma_i, \sigma_i}(G/P)$. But

$$(G/P)/(Q_1/P) \simeq H/Q \in \tau \text{ form}(\mathfrak{X} \cup \{H/Q\}) = h(\sigma_i).$$

Hence, $(G/P)/O_{\sigma_i, \sigma_i}(G/P) \in h(\sigma_i)$. Thus, for any $\sigma_j \in \sigma(G/P)$ we have $(G/P)/O_{\sigma_j, \sigma_j}(G/P) \in h(\sigma_j)$. Therefore, $G/P \in LF_\sigma(h) = \mathfrak{H}$.

Let us show that h is an integrated σ -local definition of the formation $\mathfrak{H}$. By the construction of the formation σ -function h we have

$$h(\sigma_i) = \tau \text{ form}(\mathfrak{X} \cup \{H/Q\}), \quad (1)$$

where $\mathfrak{X}$ is the set of all proper $\bar{\tau}$ -subgroups of the group H . It is clear that in equality (1) one can take as $\mathfrak{X}$ the set of all proper $\bar{\tau}$ -subgroups from G/P . But, as was established earlier, $G/P \in \mathfrak{H}$ and the formation $\mathfrak{H}$ is τ -closed. Therefore, $\mathfrak{X} \subseteq \mathfrak{H}$. Consequently, $h(\sigma_i) \subseteq \mathfrak{H}$. Now let $\sigma_j \in \sigma(G) \setminus \{\sigma_i\}$. Then $O_{\sigma_j, \sigma_j}(G)/P = O_{\sigma_j, \sigma_j}(G/P)$. Hence, since $G/O_{\sigma_j, \sigma_j}(G) \cong (G/P)/O_{\sigma_j, \sigma_j}(G/P)$ and $G/P \in \mathfrak{H}$, we have $G/O_{\sigma_j, \sigma_j}(G) \in \mathfrak{H}$. Therefore,

$$h(\sigma_j) = \tau \text{ form}(G/O_{\sigma_j, \sigma_j}(G)) \subseteq \mathfrak{H}.$$

Thus, the σ -function h is integrated.

Let $\mathfrak{M}$ be an arbitrary proper τ -closed σ -local subformation of $\mathfrak{F}$, m be its smallest τ -valued σ -local definition. Let us show that $\mathfrak{M} \subseteq \mathfrak{H}$.

Denote by f the smallest τ -valued σ -local definition of $\mathfrak{F}$. Then by lemma 5 for every $\sigma_j \in \sigma(G)$ we have $f(\sigma_j) = \tau \text{ form}(G/O_{\sigma_j, \sigma_j}(G))$ and $f(\sigma_j) = \emptyset$ for all $\sigma_j \notin \sigma(G)$. By lemma 6 we have $m \leq f$. Let $\sigma_j \in \sigma(G) \setminus \{\sigma_i\}$. Then by the construction of h we have $h(\sigma_j) = f(\sigma_j)$. Hence, for all such σ_j the inclusion $m(\sigma_j) \subseteq h(\sigma_j)$ holds. Now consider the value of f on σ_i . If $m(\sigma_i) = f(\sigma_i)$, then $G \in \mathfrak{M}$. Indeed, since $O_{\sigma_i, \sigma_i}(G) = O_{\sigma_i}(G)$ and $O_{\sigma_i}(H) = 1$, then

$$O_{\sigma_i}(G) = O_{\sigma_i}(G) \cap G = O_{\sigma_i}(G) \cap PH = P(O_{\sigma_i}(G) \cap H) = P.$$

Therefore, $G/O_{\sigma_i}(G) = G/O_{\sigma_i, \sigma_i}(G) \in f(\sigma_i) = m(\sigma_i)$. Applying now lemma 7, we conclude that $G \in \mathfrak{M}$. Hence, $\mathfrak{F} = I_{\sigma}^{\tau} \text{ form } G \subseteq \mathfrak{M}$. The latter contradicts the definition of $\mathfrak{M}$. Thus, $m(\sigma_i) \subset f(\sigma_i)$. But by lemma 8 $\tau \text{ form}(\mathfrak{X} \cup \{H/Q\}) = h(\sigma_i)$ is the unique maximal τ -subformation of the τ -irreducible formation

$$f(\sigma_i) = \tau \text{ form}(G/O_{\sigma_i, \sigma_i}(G)) = \tau \text{ form}(G/P) = \tau \text{ form } H.$$

Consequently, $m(\sigma_i) \subseteq \tau \text{ form}(\mathfrak{X} \cup \{H/Q\}) = h(\sigma_i)$. Thus, for all σ_j we have $m(\sigma_j) \subseteq h(\sigma_j)$. Hence, $\mathfrak{M} \subseteq \mathfrak{H}$.

Let us now show that $\mathfrak{H} \subseteq \mathfrak{F}$. Indeed, by lemma 5 we have

$$f(\sigma_i) = \tau \text{ form}(G/P) = \tau \text{ form } H.$$

Hence, $\mathfrak{X} \cup \{H/Q\} \subseteq f(\sigma_i)$. Therefore, $h(\sigma_i) \subseteq f(\sigma_i)$. On the other hand, by the construction of the formation σ -function h for all $\sigma_j \in \sigma(G) \setminus \{\sigma_i\}$ we have $h(\sigma_j) \subseteq f(\sigma_j)$. So, for all σ_j the inclusion $h(\sigma_j) \subseteq f(\sigma_j)$ holds. Hence, $\mathfrak{H} \subseteq \mathfrak{F}$. Assume that $\mathfrak{H} = \mathfrak{F}$. Then $t = f$, where t is the smallest τ -valued σ -local definition of $\mathfrak{H}$. Moreover, $t \leq h$. But $h(\sigma_i) \subset f(\sigma_i)$. All the more $t(\sigma_i) \subset f(\sigma_i)$. The obtained contradiction shows that $\mathfrak{H} \neq \mathfrak{F}$. The lemma is proved.

Recall that for any $\Pi \subseteq \sigma$ the symbol Π' denotes the set $\sigma \setminus \Pi$.

Lemma 24. Let $\mathfrak{M}$ be a non-empty formation, $\emptyset \neq \Pi \subseteq \sigma$. Then the canonical σ -local definition h of $\mathfrak{H} = \mathfrak{G}_{\Pi'} \mathfrak{N}_{\Pi} \mathfrak{M}$ is such that $h(\sigma_i) = \mathfrak{G}_{\sigma_i} \mathfrak{M}$ for all $\sigma_i \in \Pi$ and $h(\sigma_i) = \mathfrak{H}$ for all $\sigma_i \in \Pi'$.

Proof. Let $\mathfrak{F} = \mathfrak{G}_{\Pi'} \mathfrak{N}_{\Pi}$. By lemma 9 (1) the formation $\mathfrak{G}_{\Pi'}$ has a σ -local definition t such that $t(\sigma_i) = \mathfrak{G}_{\Pi'}$ for all $\sigma_i \in \Pi'$ and $t(\sigma_i) = \emptyset$ for all $\sigma_i \in \Pi$. Moreover, by lemma 9 (2) the formation $\mathfrak{N}_{\Pi}$ has a σ -local definition k such that $k(\sigma_i) = \mathfrak{G}_{\sigma_i}$ for all $\sigma_i \in \Pi$ and $k(\sigma_i) = \emptyset$ for all $\sigma_i \in \Pi'$. Hence, by lemma 10 (i) the formation $\mathfrak{F}$ has a σ -local definition f such that $f(\sigma_i) = \mathfrak{G}_{\Pi'} \mathfrak{N}_{\Pi}$ for all $\sigma_i \in \Pi'$ and $f(\sigma_i) = \mathfrak{G}_{\sigma_i}$ for all $\sigma_i \in \sigma$. Applying now lemma 22, we see that $\mathfrak{H}$ has a σ -local definition h such that $h(\sigma_i) = f(\sigma_i) \mathfrak{M} = \mathfrak{G}_{\Pi'} \mathfrak{N}_{\Pi} \mathfrak{M}$, if $\sigma_i \in \Pi'$, and $h(\sigma_i) = f(\sigma_i) \mathfrak{M} = \mathfrak{G}_{\sigma_i} \mathfrak{M}$, if $\sigma_i \in \Pi$. Since, obviously, $f(\sigma_i) \subseteq \mathfrak{H}$ and $h(\sigma_i) = \mathfrak{G}_{\sigma_i} h(\sigma_i)$ for all i , it follows that h is the canonical σ -local definition of $\mathfrak{H}$. The lemma is proved.

Lemma 25. Let A be a monolithic group with a non- σ -primary monolith P , $\mathfrak{X}$ be the set of all proper $\bar{\tau}$ -subgroups of the group A . Then the formation $\mathfrak{F} = I_{\sigma}^{\tau} \text{ form } A$ is I_{σ}^{τ} -irreducible, and its maximal τ -closed σ -local subformation $\mathfrak{H}$ has an integrated σ -local definition h such that $h(\sigma_i) = \tau \text{ form}(\mathfrak{X} \cup \{A/P\})$ for all $\sigma_i \in \sigma(P)$, $h(\sigma_i) = \tau \text{ form}(A/O_{\sigma_i, \sigma_i}(A))$ for all $\sigma_i \in \sigma(A) \setminus \sigma(P)$ and $h(\sigma_i) = \emptyset$ for all $\sigma_i \notin \sigma(A)$.

Proof. Let h be the formation σ -function from the condition of the lemma, and $\mathfrak{H} = LF_{\sigma}(h)$. Then by lemma 4 the formation $\mathfrak{H}$ is τ -closed. Let us show that $(\mathfrak{X} \cup \{A/P\}) \subseteq \mathfrak{H}$. Let $G \in (\mathfrak{X} \cup \{A/P\})$, $\sigma_i \in \sigma(G)$. If $\sigma_i \in \sigma(P)$, then $h(\sigma_i) = \tau \text{ form}(\mathfrak{X} \cup \{A/P\})$ and therefore $G/O_{\sigma_i, \sigma_i}(G) \in h(\sigma_i)$. Let $\sigma_i \in \sigma(G) \setminus \sigma(P)$. Then

since $G \in \mathfrak{F}$, we have $G / O_{\sigma_i, \sigma_i}(G) \in f(\sigma_i)$, where f is the smallest τ -valued σ -local definition of $\mathfrak{F}$. By lemma 5 we have $f(\sigma_i) = \tau \text{ form}(A / O_{\sigma_i, \sigma_i}(A)) = h(\sigma_i)$. Hence, $G / O_{\sigma_i, \sigma_i}(G) \in h(\sigma_i)$. Thus, for all $\sigma_i \in \sigma(G)$ we have $G / O_{\sigma_i, \sigma_i}(G) \in h(\sigma_i)$. This means that $G \in LF_\sigma(h) = \mathfrak{H}$. Consequently, $\mathfrak{X} \cup \{A / P\} \subseteq \mathfrak{H}$. From this it follows that for all $\sigma_i \in \sigma(P)$ we have $h(\sigma_i) = \tau \text{ form}(\mathfrak{X} \cup \{A / P\}) \subseteq \mathfrak{H}$. Let $\sigma_i \in \sigma(A) \setminus \sigma(P)$. Then $P \subseteq O_{\sigma_i, \sigma_i}(A)$ and therefore $A / O_{\sigma_i, \sigma_i}(A)$ is a homomorphic image of A / P . Hence, $A / O_{\sigma_i, \sigma_i}(A) \subseteq \tau \text{ form}(\mathfrak{X} \cup \{A / P\}) \subseteq \mathfrak{H}$. The latter implies $h(\sigma_i) \subseteq \mathfrak{H}$. Therefore, the σ -local definition h is integrated.

Now let $\mathfrak{M}$ be an arbitrary proper τ -closed σ -local subformation of $\mathfrak{F}$, m be its smallest τ -valued σ -local definition. Let us show that $\mathfrak{M} \subseteq \mathfrak{H}$. By lemma 6 we have $m \leq f$. By lemma 5 for all $\sigma_i \in \sigma(A) \setminus \sigma(P)$ we have $h(\sigma_i) = f(\sigma_i)$. Hence, for all such σ_i the inclusion $m(\sigma_i) \subseteq h(\sigma_i)$ holds. Let $\sigma_i \in \sigma(P)$. Then $O_{\sigma_i, \sigma_i}(A) = 1$ and by lemma 5 we obtain

$$f(\sigma_i) = \tau \text{ form}(A / O_{\sigma_i, \sigma_i}(A)) = \tau \text{ form}(A).$$

If $f(\sigma_i) = m(\sigma_i)$, then $A \in m(\sigma_i)$. But m is an integrated σ -local definition of $\mathfrak{M}$. Hence, $A \in \mathfrak{M}$ and therefore $\mathfrak{F} = I_\sigma^\tau \text{ form}(A) \subseteq \mathfrak{M} \subset \mathfrak{F}$. Hence, $\mathfrak{F} = \mathfrak{M}$. The latter contradicts the definition of $\mathfrak{M}$. Thus, $m(\sigma_i) \subset f(\sigma_i)$. By lemma 8 the formation $f(\sigma_i) = \tau \text{ form}(A)$ is τ -irreducible and every its proper τ -subformation is contained in $\tau \text{ form}(\mathfrak{X} \cup \{A / P\}) = h(\sigma_i)$. Hence, $m(\sigma_i) \subseteq h(\sigma_i)$. Thus, $m \leq h$ and therefore $\mathfrak{M} \subseteq \mathfrak{H}$.

By the construction of the σ -function h we have $h(\sigma_i) \subseteq f(\sigma_i)$ for all σ_i . Consequently, $\mathfrak{H} \subseteq \mathfrak{F}$. If $\sigma_i \in \sigma(P)$, then $h(\sigma_i) = \tau \text{ form}(\mathfrak{X} \cup \{A / P\}) \subset f(\sigma_i)$. All the more $t(\sigma_i) \subset f(\sigma_i)$, where t is the smallest τ -valued σ -local definition of $\mathfrak{H}$. The latter means that $\mathfrak{H} \subset \mathfrak{F}$. Thus, the formation $\mathfrak{F}$ is I_σ^τ -irreducible and $\mathfrak{H}$ is its unique maximal τ -closed σ -local subformation. The lemma is proved.

Lemma 26. Let $\mathfrak{M}$ be a saturated formation, $\mathfrak{H} = \mathfrak{G}_{\sigma_i} \mathfrak{G}_{\sigma_i} \mathfrak{M}$. Then $\mathfrak{F}$ is a minimal τ -closed σ -local non- $\mathfrak{H}$ -formation if and only if $\mathfrak{F} = I_\sigma^\tau \text{ form } G$, where G is a monolithic $\bar{\tau}$ -minimal non- $\mathfrak{H}$ -group with monolith $P = G^{\mathfrak{H}}$, such that $\sigma_i \in \sigma(P)$, and either P is a non- σ -primary group and G is a $\bar{\tau}$ -minimal non- $(\mathfrak{G}_{\sigma_i} \mathfrak{M})$ -group with $P = G^{\mathfrak{G}_{\sigma_i} \mathfrak{M}}$, or $G = P \rtimes H$, where $P = C_G(P)$ is a p -group, $p \in \sigma_i$, and H is a monolithic $\bar{\tau}$ -minimal non- $(\mathfrak{G}_{\sigma_i} \mathfrak{M})$ -group with monolith $Q = H^{\mathfrak{G}_{\sigma_i} \mathfrak{M}}$, such that $Q \not\subseteq \Phi(H)$ and $\sigma_i \notin \sigma(Q)$.

Proof. Let f be the smallest τ -valued σ -local definition of the formation $\mathfrak{F}$, h be the canonical σ -local definition of the formation $\mathfrak{H}$.

Necessity. By lemma 11 $\mathfrak{F}$ is a τ -closed σ -local formation generated by a monolithic $\bar{\tau}$ -minimal non- $\mathfrak{H}$ -group G with monolith $P = G^{\mathfrak{H}}$, such that the formation $f(\sigma_j)$ is a minimal τ -closed non- $h(\sigma_j)$ -formation for all $\sigma_j \in \sigma(P)$. Since $\mathfrak{H} = \mathfrak{G}_{\sigma_i} \mathfrak{H}$, we have $\sigma_i \in \sigma(P)$.

Let P be a non- σ -primary group. Then by lemma 25 the formation $\mathfrak{F}$ is I_σ^τ -irreducible and its maximal τ -closed σ -local subformation $\mathfrak{M}_1$ has an integrated τ -valued σ -local definition m such that $m(\sigma_j) = \tau \text{ form}(\mathfrak{X} \cup \{G / P\})$, where $\mathfrak{X}$ is the set of all proper $\bar{\tau}$ -subgroups of the group G for all $\sigma_j \in \sigma(P)$. But by condition $\mathfrak{M}_1 \subseteq \mathfrak{H}$. Hence, $m \leq h$. By lemma 24 we have $h(\sigma_i) = \mathfrak{G}_{\sigma_i} \mathfrak{M}$. Thus, $(\mathfrak{X} \cup \{G / P\}) \subseteq \mathfrak{G}_{\sigma_i} \mathfrak{M}$. It is clear that G does not belong to $\mathfrak{G}_{\sigma_i} \mathfrak{M}$. Hence, G is a $\bar{\tau}$ -minimal non- $(\mathfrak{G}_{\sigma_i} \mathfrak{M})$ -group and $G / P \in \mathfrak{G}_{\sigma_i} \mathfrak{M}$, i. e. $P = G^{\mathfrak{G}_{\sigma_i} \mathfrak{M}}$.

Let P be a σ -primary group. Then P is a σ_i -group. And let H be a group of minimal order from $f(\sigma_i) \setminus h(\sigma_i)$. Since the formation $f(\sigma_i)$ is τ -closed, H is a monolithic $\bar{\tau}$ -minimal non- $h(\sigma_i)$ -group. Let Q be the monolith of H . Then $Q = H^{h(\sigma_i)} = H^{\mathfrak{G}_{\sigma_i} \mathfrak{M}}$. Let us show that Q is not contained in $\Phi(H)$. Suppose the contrary. Then since $H / Q \in h(\sigma_i)$, we have $H / \Phi(H) \in h(\sigma_i)$. By the condition of the lemma the formation $\mathfrak{M}$ is saturated, and hence local. Therefore, by lemma 12 the formation $h(\sigma_i) = \mathfrak{G}_{\sigma_i} \mathfrak{M}$ is local, and hence $h(\sigma_i)$ is saturated. Consequently, $H \in h(\sigma_i)$. The obtained contradiction shows that $\Phi(H) = 1$.

Since $h(\sigma_i) = \mathfrak{G}_{\sigma_i} \mathfrak{M}$, we have $O_{\sigma_i}(H) = 1$. Consequently, by lemma 13 there exists a faithful irreducible $F_p H$ -module P , where $p \in \sigma_i$. Let $F = P \rtimes H$. Then $P = C_F(P)$ is the minimal normal σ_i -subgroup of F . Since the σ -local definition f is integrated, $F \in \mathfrak{F}$ by lemma 7. It is clear that $O_{\sigma_i, \sigma_i}(F) = O_{\sigma_i}(F)$. Moreover, since $O_{\sigma_i}(H) = 1$, we have

$$O_{\sigma_i}(F) = O_{\sigma_i}(F) \cap F = O_{\sigma_i}(F) \cap PH = P(O_{\sigma_i}(F) \cap H) = P.$$

Denote by h_1 the smallest τ -valued σ -local definition of $\mathfrak{H}_1 = l_\sigma^\tau \text{ form } F$. By lemma 5 we have $h_1(\sigma_i) = \tau \text{ form}(F / O_{\sigma_i, \sigma_i}(F)) = \tau \text{ form}(F / P) = \tau \text{ form } H$. If $\mathfrak{H}_1 \subset \mathfrak{F}$, then by condition $\mathfrak{H}_1 \subseteq \mathfrak{H}$. Hence, by lemma 6 we have the embedding $h_1 \leq h$. The latter contradicts the definition of the group H . Thus, $\mathfrak{H}_1 = \mathfrak{F}$, i. e. $\mathfrak{F} = l_\sigma^\tau \text{ form } F$.

Suppose that $\sigma_i \in \sigma(Q)$. Since $O_{\sigma_i}(H) = 1$, Q is a non- σ -primary group. It is clear that $H \in \mathfrak{F}$ and $O_{\sigma_i, \sigma_i}(H) = 1$. Suppose that the formation $\mathfrak{H}_2 = l_\sigma^\tau \text{ form } H \neq \mathfrak{F}$. Then by condition $\mathfrak{H}_2 \subseteq \mathfrak{H}$. Hence, $H \simeq H / O_{\sigma_i, \sigma_i}(H) \in h(\sigma_i)$, which contradicts the definition of the group H . Consequently, $\mathfrak{H}_2 = \mathfrak{F}$. Moreover, repeating the above arguments, we obtain that H is a minimal non- $(\mathfrak{G}_{\sigma_i} \mathfrak{M})$ -group and $Q = H^{\mathfrak{G}_{\sigma_i} \mathfrak{M}}$. Thus, we arrive at the situation already considered. Therefore, $\sigma_i \notin \sigma(Q)$.

Sufficiency. By lemma 11, to prove that $\mathfrak{F}$ is a minimal τ -closed σ -local non- $\mathfrak{H}$ -formation, it suffices to establish that $f(\sigma_i)$ is a minimal τ -closed non- $h(\sigma_i)$ -formation for all $\sigma_i \in \sigma(P)$.

Let P be a non- σ -primary group. Then $O_{\sigma_i, \sigma_i}(G) = 1$ for all $\sigma_i \in \sigma(P)$. Consequently, by lemma 5 for all such σ_i we have $f(\sigma_i) = \tau \text{ form } G$. By lemma 8 the formation $f(\sigma_i)$ is τ -irreducible and its maximal τ -closed subformation coincides with $\tau \text{ form}(\mathfrak{X} \cup \{G/P\})$, where $\mathfrak{X}$ is the set of all proper τ -subgroups of G . By condition G is a τ -minimal non- $(\mathfrak{G}_{\sigma_i} \mathfrak{M})$ -group and $P = G^{\mathfrak{G}_{\sigma_i} \mathfrak{M}}$. Hence, if $\mathfrak{X}_1 = S_\tau(G/P)$, then $\mathfrak{X}_1 \subseteq h(\sigma_i) = \mathfrak{G}_{\sigma_i} \mathfrak{M}$. Therefore, by lemmas 14 and 24 we have

$$\tau \text{ form}(\mathfrak{X} \cup \{G/P\}) = \tau \text{ form}(\mathfrak{X} \cup \mathfrak{X}_1) = \text{form}(\mathfrak{X} \cup \mathfrak{X}_1) \subseteq h(\sigma_i) = \mathfrak{G}_{\sigma_i} \mathfrak{M} \subseteq \mathfrak{H} = h(\sigma_i)$$

for all $\sigma_i \in \sigma(P) \setminus \{\sigma_i\}$. Since also $P = G^\mathfrak{H} = G^{\mathfrak{G}_{\sigma_i} \mathfrak{M}}$, it follows that $f(\sigma_i)$ is not contained in $h(\sigma_i)$ for all $\sigma_i \in \sigma(P)$. Thus, if P is a non- σ -primary group, then $f(\sigma_i)$ is a minimal τ -closed non- $h(\sigma_i)$ -formation for all $\sigma_i \in \sigma(P)$.

Let P be a σ_i -group. Then $O_{\sigma_i, \sigma_i}(G) = O_{\sigma_i}(G)$. Moreover, since $O_{\sigma_i}(H) = 1$, we have

$$O_{\sigma_i}(G) = O_{\sigma_i}(G) \cap G = O_{\sigma_i}(G) \cap PH = P(O_{\sigma_i}(G) \cap H) = P.$$

By lemma 5 we have

$$f(\sigma_i) = \tau \text{ form}(G / O_{\sigma_i, \sigma_i}(G)) = \tau \text{ form } H.$$

Since H is monolithic and $\Phi(H) = 1$, by lemma 8 $\tau \text{ form}(\mathfrak{X} \cup \{H/Q\})$ is the unique maximal τ -closed subformation of $f(\sigma_i)$, where $\mathfrak{X}$ is the set of all proper τ -subgroups of the group H . Now, as above, one can show that

$$\tau \text{ form}(\mathfrak{X} \cup \{H/Q\}) \subseteq h(\sigma_i) = \mathfrak{G}_{\sigma_i} \mathfrak{M}.$$

Thus, $f(\sigma_i)$ is a minimal τ -closed non- $h(\sigma_i)$ -formation. Hence, $\mathfrak{F}$ is a minimal τ -closed σ -local non- $\mathfrak{H}$ -formation by lemma 11. The lemma is proved.

Lemma 27. Let $\mathfrak{M}$ be a non-empty Abelian formation, $\mathfrak{H} = \mathfrak{G}_{\sigma_i} \mathfrak{G}_{\sigma_i} \mathfrak{M}$. Then $\mathfrak{F}$ is a minimal τ -closed σ -local non- $\mathfrak{H}$ -formation if and only if $\mathfrak{F} = l_\sigma^\tau \text{ form } G$, where G is a monolithic τ -minimal non- $\mathfrak{H}$ -group with monolith $P = G^\mathfrak{H}$, such that $\sigma_i \in \sigma(P)$, and either $P = G^{\mathfrak{G}_{\sigma_i} \mathfrak{M}}$ is a non- σ -primary group and G is a τ -minimal non- $(\mathfrak{G}_{\sigma_i} \mathfrak{M})$ -group, or $G = P \rtimes H$, where $P = C_G(P)$ is a p -group, $p \in \sigma_i$, and H satisfies one of the following conditions:

- 1) H is a monolithic τ -minimal non- $(\mathfrak{G}_{\sigma_i} \mathfrak{M})$ -group with monolith $Q = H^{\mathfrak{G}_{\sigma_i} \mathfrak{M}} \not\subseteq \Phi(H)$, such that $\sigma_i \notin \sigma(Q)$;
- 2) H is a minimal non- $\mathfrak{M}$ -group of one of the following types:
 - a) the quaternion group of order 8, if $2 \notin \sigma_i$;
 - b) an extraspecial group of order q^3 of prime odd exponent $q \notin \sigma_i$;
 - c) a cyclic q -group, $q \notin \sigma_i$.

Proof. Let f be the smallest τ -valued σ -local definition of $\mathfrak{F}$, h be the canonical σ -local definition of $\mathfrak{H}$.

Necessity. Let $\mathfrak{F}$ be an $\mathfrak{H}_\sigma^\tau$ -critical formation. Suppose that $\mathfrak{F}$ is not contained in $(\mathfrak{G}_{\sigma_i} \mathfrak{G}_{\sigma_i}) \mathfrak{N} = \mathfrak{H}_1$, where $\mathfrak{N}$ is the class of all nilpotent groups. Since $\mathfrak{M} \subseteq \mathfrak{N}$, we have $(\mathfrak{G}_{\sigma_i} \mathfrak{G}_{\sigma_i}) \mathfrak{M} \subseteq (\mathfrak{G}_{\sigma_i} \mathfrak{G}_{\sigma_i}) \mathfrak{N} = \mathfrak{H}_1$. Hence, in this case $\mathfrak{F}$ is a minimal τ -closed σ -local non- $\mathfrak{H}_1$ -formation. Consequently, by lemma 26 we have that $\mathfrak{F}$ is a τ -closed σ -local formation generated by a monolithic τ -minimal non- $\mathfrak{H}_1$ -group G with monolith $P = G^{\mathfrak{H}_1}$, such that $\sigma_i \in \sigma(P)$, and either P is a non- σ -primary group and G is a τ -minimal non- $(\mathfrak{G}_{\sigma_i} \mathfrak{N})$ -group, and $P = G^{\mathfrak{G}_{\sigma_i} \mathfrak{N}}$, or $G = P \rtimes H$, where $P = C_G(P)$ is a p -group, $p \in \sigma_i$, and H is a monolithic τ -minimal non- $(\mathfrak{G}_{\sigma_i} \mathfrak{N})$ -group with monolith $Q = H^{\mathfrak{G}_{\sigma_i} \mathfrak{N}}$, such that $Q \not\subseteq \Phi(H)$ and $\sigma_i \notin \sigma(Q)$.

Let P be a non- σ -primary group. Then by lemma 25 the unique maximal τ -closed σ -local subformation $\mathfrak{M}_1$ of $\mathfrak{F}$ has an integrated τ -valued σ -local definition m such that $m(\sigma_i) = \tau \text{ form}(\mathfrak{X} \cup \{G/P\})$, where $\mathfrak{X}$ is the set of all proper $\bar{\tau}$ -subgroups of the group G . By lemma 1 (4), (5) we have $m \leq t$, where t is the canonical σ -local definition of $\mathfrak{M}_1$. Since by condition $\mathfrak{M}_1 \subseteq \mathfrak{H}$, by lemma 15 the embedding $m \leq t \leq h$ holds. By lemma 24 we have $h(\sigma_i) = \mathfrak{G}_{\sigma_i} \mathfrak{M}$. Hence, $(\mathfrak{X} \cup \{G/P\}) \subseteq \mathfrak{G}_{\sigma_i} \mathfrak{M}$. Thus, if P is a non- σ -primary group, then $P = G^{\mathfrak{G}_{\sigma_i} \mathfrak{M}}$ and G is a $\bar{\tau}$ -minimal non- $(\mathfrak{G}_{\sigma_i} \mathfrak{M})$ -group.

Let P be a σ -primary group, $\mathfrak{X}$ be the set of all proper $\bar{\tau}$ -subgroups of H . Then, taking into account lemma 23 and reasoning as in the previous paragraph, we obtain that $\mathfrak{X} \cup \{H/Q\} \subseteq \mathfrak{G}_{\sigma_i} \mathfrak{M}$. Hence, H is a $\bar{\tau}$ -minimal non- $(\mathfrak{G}_{\sigma_i} \mathfrak{M})$ -group. Since also $H/Q \in \mathfrak{G}_{\sigma_i} \mathfrak{M}$, we have $Q = H^{\mathfrak{G}_{\sigma_i} \mathfrak{M}}$. Thus, H satisfies condition 1).

Note that in both considered cases we have $P = G^{\mathfrak{H}}$. Indeed, if P is non- σ -primary, then since $G/P \in \mathfrak{G}_{\sigma_i} \mathfrak{M} \subseteq \mathfrak{H}$ we have $G/P \in \mathfrak{H}$. If P is σ -primary, then since $Q = H^{\mathfrak{G}_{\sigma_i} \mathfrak{M}}$ and Q is a σ_i -group, we again obtain $G/P \simeq H \in \mathfrak{G}_{\sigma_i} \mathfrak{G}_{\sigma_i} \mathfrak{M} = \mathfrak{H}$. Assume that $G \in \mathfrak{H}$. Then $G \in \mathfrak{H}_1$, but $P = G^{\mathfrak{H}}$. Contradiction. Consequently, $P = G^{\mathfrak{H}}$.

Thus, $\mathfrak{F} = I_{\sigma}^{\tau} \text{ form } G$, where G is a group satisfying the condition of the lemma.

Consider the case when $\mathfrak{F} \subseteq \mathfrak{H}_1$. By lemma 11 $\mathfrak{F}$ is a τ -closed σ -local formation generated by a monolithic $\bar{\tau}$ -minimal non- $\mathfrak{H}$ -group F with monolith $F^{\mathfrak{H}}$, such that for all $\sigma_j \in \sigma(F^{\mathfrak{H}})$ it holds that $f(\sigma_j)$ is a minimal τ -closed non- $h(\sigma_j)$ -formation. Since $F/F^{\mathfrak{H}} \in \mathfrak{H} = \mathfrak{G}_{\sigma_i} \mathfrak{G}_{\sigma_i} \mathfrak{M}$, we have $\sigma_i \in \sigma(F^{\mathfrak{H}})$. By lemma 24 the formation $\mathfrak{G}_{\sigma_i} \mathfrak{M}$ is the value at σ_i of the canonical σ -local definition of $\mathfrak{H}_1$. Since by our assumption $\mathfrak{F} \subseteq \mathfrak{H}_1$, we have $f(\sigma_i) \subseteq \mathfrak{G}_{\sigma_i} \mathfrak{M}$. By lemma 5 (3) we have $f(\sigma_i) = \tau \text{ form}(A \mid A \in f(\sigma_i), O_{\sigma_i}(A) = 1)$. Hence, $f(\sigma_i) \subseteq \mathfrak{N}_{\sigma_i}$, where $\mathfrak{N}_{\sigma_i}$ is the formation of all nilpotent σ_i -groups.

By lemma 16 all proper subformations of $f(\sigma_i)$ are hereditary. From this and from the fact that every proper τ -closed subformation of $f(\sigma_i)$ is contained in $h(\sigma_i) = \mathfrak{G}_{\sigma_i} \mathfrak{M}$, it follows that every proper subformation of $f(\sigma_i)$ is contained in $\mathfrak{M}$. On the other hand, it is clear that $f(\sigma_i)$ is not contained in $\mathfrak{M}$. Hence, $f(\sigma_i)$ is a minimal τ -closed non- $\mathfrak{M}$ -formation.

Let $f(\sigma_i)$ not be contained in $\mathfrak{A}$. Then since $\mathfrak{M} \subseteq \mathfrak{A}$, it follows that $f(\sigma_i)$ is a minimal non-Abelian formation. By lemma 17 $f(\sigma_i)$ is the formation generated by H , where H is either the quaternion group of order 8 or an extraspecial group of order q^3 of prime odd exponent q . Since $f(\sigma_i) \subseteq \mathfrak{N}_{\sigma_i}$, in the first case $2 \notin \sigma_i$, and in the second case $q \notin \sigma_i$.

Since in both these cases the group H is monolithic and $O_p(H) = 1$, where $p \in \sigma_i$, by lemma 13 there exists a faithful irreducible $F_p H$ -module P . Let $G = P \rtimes H$. Then

$$O_{\sigma_i, \sigma_i}(G) = O_{\sigma_i}(G) = O_{\sigma_i}(G) \cap PH = P(O_{\sigma_i}(G) \cap H) = P,$$

since H is a σ_i -group. Hence $G/O_{\sigma_i}(G) \simeq H \in f(\sigma_i)$. Therefore, $G \in \mathfrak{F}$ by lemma 7. If the τ -closed σ -local formation $\mathfrak{F}_1$, generated by the group G , is different from $\mathfrak{F}$, then by condition $\mathfrak{F}_1 \subseteq \mathfrak{H}$. But then $G/O_{\sigma_i, \sigma_i}(G) \simeq H \in h(\sigma_i)$. Consequently, $f(\sigma_i) \subseteq h(\sigma_i)$. Contradiction. Hence, $\mathfrak{F}_1 = \mathfrak{F}$. Moreover, obviously, $P = C_G(P) = G^{\mathfrak{H}}$. Since $f(\sigma_i)$ is nilpotent, we have $\text{form } \mathfrak{X} \subseteq \text{form } H = f(\sigma_i)$, where $\mathfrak{X}$ is the set of all proper subgroups of H . Since all proper subgroups of H are Abelian and the formation $f(\sigma_i)$ is non-Abelian, we have $\text{form } \mathfrak{X} \subset f(\sigma_i)$. Hence, since $f(\sigma_i)$ is a minimal non- $\mathfrak{M}$ -formation, H is a minimal non- $\mathfrak{M}$ -group.

Let $f(\sigma_i) \subseteq \mathfrak{A}$ and H be a group of minimal order from $f(\sigma_i) \setminus \mathfrak{M}$. Since $f(\sigma_i)$ is hereditary, H is a minimal non- $\mathfrak{M}$ -group. Moreover, because $f(\sigma_i)$ is a minimal non- $\mathfrak{M}$ -formation, $f(\sigma_i)$ is the formation generated by the group H . Since the formation $f(\sigma_i)$ is Abelian, H is a cyclic q -group, where $q \in \sigma_j \in \sigma \setminus \{\sigma_i\}$.

Considering that $f(\sigma_i) \subseteq \mathfrak{N}_{\sigma_i}$, then $O_p(H) = 1$, where $p \in \sigma_i$. Therefore, by lemma 13 there exists a simple faithful $F_p H$ -module P . Let $G = P \rtimes H$. Since H does not belong to $\mathfrak{M}$, then, as above, one can show that $\mathfrak{F} = I_{\sigma}^{\tau} \text{ form } G$. Moreover, it is clear that $P = C_G(P) = G^{\mathfrak{H}}$ and H satisfy condition 2).

Sufficiency. By lemma 11, to prove that $\mathfrak{F}$ is a minimal τ -closed σ -local non- $\mathfrak{H}$ -formation, it suffices to show that $f(\sigma_r)$ is a minimal τ -closed non- $h(\sigma_r)$ -formation for all $\sigma_r \in \sigma(P)$.

Let P be non- σ -primary. Then $O_{\sigma_r, \sigma_r}(G) = 1$ for every $\sigma_r \in \sigma(P)$. Consequently, $f(\sigma_r) = \tau \text{ form } G$. By lemma 8 the formation $f(\sigma_r)$ has a unique maximal τ -closed subformation, coinciding with $\tau \text{ form}(\mathfrak{X} \cup \{G/P\})$, where $\mathfrak{X}$ is the set of all proper τ -subgroups of the group G . But by condition G is a $\bar{\tau}$ -minimal non- $(\mathfrak{G}_{\sigma_i} \mathfrak{M})$ -group

and $P = G^{\mathfrak{G}_{\sigma_i} \mathfrak{M}}$. Hence, $\mathfrak{X} \cup \{G/P\} \subseteq \mathfrak{G}_{\sigma_i} \mathfrak{M} = h(\sigma_i)$. Consequently, the inclusion $\text{form}(\mathfrak{X} \cup \{G/P\}) \subseteq h(\sigma_i)$ holds. But $\text{form}(\mathfrak{X} \cup \{G/P\}) = \tau \text{form}(\mathfrak{X} \cup \{G/P\})$. Indeed, if $L/P \in \tau(G/P)$ and $L \neq G$, then $L \in \mathfrak{X}$. Therefore, $L/P \in H(\mathfrak{X} \cup \{G/P\})$. Hence,

$$S_{\tau}(\mathfrak{X} \cup \{G/P\}) \subseteq H(\mathfrak{X} \cup \{G/P\}).$$

But by lemma 18 we have

$$\begin{aligned} \tau \text{form}(\mathfrak{X} \cup \{G/P\}) &= HR_0 S_{\tau}(\mathfrak{X} \cup \{G/P\}) = \text{form}(S_{\tau}(\mathfrak{X} \cup \{G/P\})) = \\ &= \text{form}(H(\mathfrak{X} \cup \{G/P\})) = \text{form}(\mathfrak{X} \cup \{G/P\}). \end{aligned}$$

Thus,

$$\tau \text{form}(\mathfrak{X} \cup \{G/P\}) \subseteq h(\sigma_i).$$

On the other hand, since by lemma 23 for any $\sigma_i \in \sigma \setminus \{\sigma_i\}$ we have $h(\sigma_i) = \mathfrak{G}_{\sigma_i} \mathfrak{G}_{\sigma_i} \mathfrak{M}$, it follows that for all such σ_i the inclusion $\tau \text{form}(\mathfrak{X} \cup \{G/P\}) \subseteq h(\sigma_i)$ also holds. Since $P = G^{\mathfrak{H}}$, it follows that $f(\sigma_i)$ is not contained in $h(\sigma_i)$ for all $\sigma_i \in \sigma(P)$. Thus, if P is a non- σ -primary group, then $f(\sigma_i)$ is a minimal τ -closed non- $h(\sigma_i)$ -formation for all $\sigma_i \in \sigma(G^{\mathfrak{H}})$. Consequently, $\mathfrak{F}$ is a minimal τ -closed σ -local non- $\mathfrak{H}$ -formation.

Now let P be a σ_i -group. Then $O_{\sigma_i, \sigma_i}(G) = O_{\sigma_i}(G)$ and since $O_{\sigma_i}(H) = 1$ we have

$$O_{\sigma_i}(G) = O_{\sigma_i}(G) \cap G = O_{\sigma_i}(G) \cap PH = P(O_{\sigma_i}(G) \cap H) = P.$$

Consequently, $f(\sigma_i) = \tau \text{form}(G/O_{\sigma_i, \sigma_i}(G)) = \tau \text{form} H$ by lemma 5. It is clear that $f(\sigma_i)$ is not contained in $h(\sigma_i)$.

Let H satisfy condition 1). Then by lemma 8 the formation $f(\sigma_i)$ has a unique maximal τ -closed subformation $\mathfrak{M}_1 = \tau \text{form}(\mathfrak{X} \cup \{H/Q\})$, where $\mathfrak{X}$ is the set of all proper τ -subgroups of H . By condition H is a τ -minimal non- $(\mathfrak{G}_{\sigma_i} \mathfrak{M})$ -group and $Q = H^{\mathfrak{G}_{\sigma_i} \mathfrak{M}}$. Hence,

$$\tau \text{form}(\mathfrak{X} \cup \{H/Q\}) = \text{form}(\mathfrak{X} \cup \{H/Q\}) \subseteq h(\sigma_i) = \mathfrak{G}_{\sigma_i} \mathfrak{M}.$$

It is clear that $f(\sigma_i)$ is not contained in $h(\sigma_i)$. Thus, $f(\sigma_i)$ is a minimal τ -closed non- $h(\sigma_i)$ -formation, and hence $\mathfrak{F}$ is an $\mathfrak{H}_{\sigma}^{\tau}$ -critical formation.

Now let H satisfy condition 2) and let H be the quaternion group of order 8, $2 \notin \sigma_i$. And let $\mathfrak{M}_1$ be the formation of Abelian groups of exponent dividing 4. Suppose $\mathfrak{M}_1 \not\subseteq f(\sigma_i)$ and let K be a group of minimal order from $\mathfrak{M}_1 \setminus f(\sigma_i)$. Then K is monolithic. Consequently, K is a cyclic primary group. Thus, either $|K| = 4$, or $|K| = 2$. But the quaternion group of order 8 has cyclic subgroups of orders 2 and 4. Consequently, $K \in f(\sigma_i)$. The obtained contradiction shows that $\mathfrak{M}_1 \subseteq f(\sigma_i)$. Let $\mathfrak{H}_1$ be an arbitrary proper subformation of $f(\sigma_i)$. Then by lemma 17 we have $\mathfrak{H}_1 \subseteq \mathfrak{A}$. Assume that $\mathfrak{H}_1 \not\subseteq \mathfrak{M}_1$. Let M be a group of minimal order from $\mathfrak{H}_1 \setminus \mathfrak{M}_1$. Then M is monolithic. Since the group M is Abelian, it is cyclic. By lemma 19 the exponent of M divides the exponent of H . Thus, the exponent of M is 2 or 4. But then, obviously, $M \in \mathfrak{M}_1$. The obtained contradiction shows that every proper subformation of $f(\sigma_i)$ is contained in $\mathfrak{M}_1$. But, obviously, $\mathfrak{M}_1 = \text{form}(Z_4)$, where Z_4 is the cyclic group of order 4. Hence, $\mathfrak{M}_1 \subseteq \tau \text{form} \mathfrak{X}$, where $\mathfrak{X}$ is the set of all proper subgroups of H . Thus,

$$\mathfrak{M}_1 = \text{form}(Z_4) = \tau \text{form} \mathfrak{X}.$$

Since by condition H is a minimal non- $\mathfrak{M}$ -group, we have $\mathfrak{M}_1 \subseteq \mathfrak{M}$. Consequently, $f(\sigma_i)$ is a minimal τ -closed non- $h(\sigma_i)$ -formation.

Now let H be an extraspecial group of order q^3 of prime odd exponent $q \notin \sigma_i$. Then by lemma 17 $\tau \text{form} H = f(\sigma_i)$ is a minimal τ -closed non-Abelian formation. Let M be a maximal subformation of $f(\sigma_i)$. Then M is an Abelian formation and every non-identity group from M has exponent q . Since K is a minimal non- $h(\sigma_i)$ -group, $h(\sigma_i)$ contains groups of order q . But then $M \subseteq h(\sigma_i)$. Consequently, $f(\sigma_i)$ is a minimal non- $h(\sigma_i)$ -formation.

Finally, let H be a cyclic q -group, $q \notin \sigma_i$. Let $|H| = q^n$ and M be a maximal subgroup of H . Then $M \in h(\sigma_i)$, but $H \notin h(\sigma_i)$. Hence, $f(\sigma_i) \not\subseteq h(\sigma_i)$ and $\text{form} M \subseteq h(\sigma_i)$. Let us show that every proper subformation of

$f(\sigma_i)$ is contained in form M . Suppose that in $f(\sigma_i)$ there is a proper subformation $\mathfrak{M}$ such that $\mathfrak{M} \not\subseteq \text{form } M$, and let A be a group of minimal order from $\mathfrak{M} \setminus \text{form } M$. Then A is monolithic. Since $f(\sigma_i)$ is Abelian, A is a cyclic group of order q^m . According to lemma 19 the exponent q^m of A divides the exponent q^n of H . If $m = n$, then $A \simeq H$ and then $f(\sigma_i) \subseteq \text{form } A \subseteq \mathfrak{M} \subset f(\sigma_i)$. The obtained contradiction shows that $m \neq n$. But the order of M is q^{n-1} . Hence, either A is isomorphic to some proper subgroup of M , or $A \simeq M$. This means that $A \in \text{form } M$. Contradiction. Consequently, $f(\sigma_i)$ is a minimal non- $h(\sigma_i)$ -formation.

Thus, in all considered cases $f(\sigma_i)$ is a minimal τ -closed non- $h(\sigma_i)$ -formation. Therefore, $\mathfrak{F}$ is an $\mathfrak{H}_\sigma^\tau$ -critical formation by lemma 11. The lemma is proved.

Lemma 28. Let h be the canonical definition of $\mathfrak{H}$ and $G = P \rtimes K$ be a monolithic group with monolith $P = C_G(P) = G^\mathfrak{H} - a$ p -group, $p \in \sigma_i \in \sigma$, $O_{\sigma_i}(K) = 1$. Then G is a $\bar{\tau}$ -minimal non- $\mathfrak{H}$ -group if and only if K is a $\bar{\tau}$ -minimal non- $h(\sigma_i)$ -group.

Proof. Sufficiency. Let K be a $\bar{\tau}$ -minimal non- $h(\sigma_i)$ -group and $M \in \bar{\tau}(G) \setminus \{G\}$. Let us show that $M \in \mathfrak{H}$. If $G = PM$, then, obviously, $M \cap P = 1$. Hence, $M \simeq M/P \cap M \simeq MP/P = G/P \in \mathfrak{H}$. Let $PM \neq G$. Then $PM/P \in \bar{\tau}(G/P) \setminus \{G/P\}$. But $G/P \simeq K$. Consequently, in K there exists a proper $\bar{\tau}$ -subgroup D such that $MP/P \simeq D \in h(\sigma_i)$. So,

$$M/M \cap P \simeq MP/P \in h(\sigma_i) = \mathfrak{G}_{\sigma_i} h(\sigma_i).$$

Hence, $M \in h(\sigma_i)$. Thus, every proper $\bar{\tau}$ -subgroup of G belongs to $\mathfrak{H}$. Since by condition $P = G^\mathfrak{H}$, it follows that G is a $\bar{\tau}$ -minimal non- $\mathfrak{H}$ -group.

Necessity. Let G be a $\bar{\tau}$ -minimal non- $\mathfrak{H}$ -group and $M \in \bar{\tau}(K) \setminus \{K\}$. If under the isomorphism $G/P \simeq K$ the subgroup M corresponds to the subgroup D/P in the group G/P , then $D \in \bar{\tau}(G) \setminus \{G\}$. Consequently, $D \in \mathfrak{H}$. Therefore, $D/O_{\sigma_i, \sigma_i}(D) \in h(\sigma_i)$. Then since

$$D = D \cap G = D \cap PK = P \rtimes (D \cap K), \quad P = C_G(P) \text{ and } O_{\sigma_i}(K) = 1,$$

we have $O_{\sigma_i, \sigma_i}(D) = O_{\sigma_i}(D)$. It is also clear that $O_{\sigma_i}(D)/P = O_{\sigma_i}(D/P)$. Consequently, since $D/O_{\sigma_i, \sigma_i}(D) \in h(\sigma_i)$, it follows that $M/O_{\sigma_i}(M) \in h(\sigma_i) = \mathfrak{G}_{\sigma_i} h(\sigma_i)$. Hence, $M \in h(\sigma_i)$. Thus, every proper $\bar{\tau}$ -subgroup of K belongs to $h(\sigma_i)$. If, moreover, $K \in h(\sigma_i)$, then $G \in \mathfrak{H}$ by lemma 7, which is impossible. Therefore, K is a $\bar{\tau}$ -minimal non- $h(\sigma_i)$ -group. The lemma is proved.

Proof of theorem 1

Necessity. According to the condition, the formation $\mathfrak{H}$ has a σ -local definition f all of whose non-Abelian values are σ -local. By lemma 20 we can assume that such a σ -local definition is integrated. By lemma 1 (2) we have

$$\mathfrak{H} = \left(\bigcap_{\sigma_i \in \sigma(\mathfrak{H})} \mathfrak{G}_{\sigma_i'} \mathfrak{G}_{\sigma_i} f(\sigma_i) \right) \cap \mathfrak{G}_{\sigma(\mathfrak{H})}.$$

Since by condition every proper τ -closed σ -local subformation of $\mathfrak{F}$ is contained in $\mathfrak{H}$, but $\mathfrak{F} \not\subseteq \mathfrak{H}$, then either $\mathfrak{F}$ is a minimal τ -closed σ -local non- $\mathfrak{G}_{\sigma(\mathfrak{H})}$ -formation, or there exists $\sigma_i \in \sigma(\mathfrak{H})$ such that $\mathfrak{F}$ is a minimal τ -closed σ -local non- $(\mathfrak{G}_{\sigma_i'} \mathfrak{G}_{\sigma_i} f(\sigma_i))$ -formation.

Consider first the case when $\mathfrak{F}$ is a minimal τ -closed σ -local non- $\mathfrak{G}_{\sigma(\mathfrak{H})}$ -formation. Then by lemma 21 we have $\mathfrak{F} = \mathfrak{G}_{\sigma_i} = I_\sigma^\tau \text{ form } G$, where G is a simple σ_i -group such that $\sigma_i \notin \sigma(\mathfrak{H})$ and $\tau(G) \subseteq \{1, G\}$. It is clear that G is a $\bar{\tau}$ -minimal non- $\mathfrak{H}$ -group and $G = G^\mathfrak{H}$. Thus, the group G satisfies condition 1) of the theorem.

Now let $\sigma_i \in \sigma(\mathfrak{H})$ be such that $\mathfrak{F}$ is a τ -closed σ -local non- $(\mathfrak{G}_{\sigma_i'} \mathfrak{G}_{\sigma_i} f(\sigma_i))$ -formation.

Assume first that $f(\sigma_i)$ is a σ -local formation. By lemma 1 (5) $h(\sigma_j) = \mathfrak{G}_{\sigma_j} f(\sigma_j)$ for any σ_j . Since by lemma 1 (2) the formation $f(\sigma_i)$ is saturated, by lemma 26 we have $\mathfrak{F} = I_\sigma^\tau \text{ form } G$, where G is a monolithic group with monolith $P = G^{\mathfrak{G}_{\sigma_i'} h(\sigma_i)}$, such that $\sigma_i \in \sigma(P)$ and either P is a non- σ -primary group, G is a $\bar{\tau}$ -minimal non- $h(\sigma_i)$ -group and $P = G^{h(\sigma_i)}$, or $G = P \rtimes K$, where $P = C_G(P)$ is a p -group, $p \in \sigma_i$, and K is a monolithic $\bar{\tau}$ -minimal non- $h(\sigma_i)$ -group with monolith $Q = K^{h(\sigma_i)}$, such that $Q \not\subseteq \Phi(K)$ and $\sigma_i \notin \sigma(Q)$.

Denote by $\mathfrak{M}$ the maximal τ -closed σ -local subformation of $\mathfrak{F}$. Consider the case when P is a non- σ -primary group. By lemma 25 the formation $\mathfrak{M}$ has an integrated σ -local definition m such that $m(\sigma_j) = \tau \text{ form}(\mathfrak{X} \cup \{G/P\})$ for all $\sigma_j \in \sigma(P)$, where $\mathfrak{X}$ is the set of all proper $\bar{\tau}$ -subgroups of G . By lemma 15 we have $m \leq h$. Thus, for all $\sigma_j \in \sigma(P)$ it holds that all proper $\bar{\tau}$ -subgroups of G belong to $h(\sigma_j)$ and $G/P \in h(\sigma_j)$. Suppose that there exists $\sigma_j \in \sigma(P)$ such that $G \in h(\sigma_j)$. Then since the σ -local definition h is integrated, we have

$$G \in h(\sigma_j) \subseteq \mathfrak{H} \subseteq \mathfrak{G}_{\sigma_i} \mathfrak{G}_{\sigma_i} f(\sigma_i).$$

But $P = G^{\mathfrak{G}_{\sigma_i} h(\sigma_i)} = G^{\mathfrak{G}_{\sigma_i} \mathfrak{G}_{\sigma_i} f(\sigma_i)}$. The obtained contradiction shows that G is a $\bar{\tau}$ -minimal non- $h(\sigma_j)$ -group for all $\sigma_j \in \sigma(P)$. It is clear that $P = G^{\mathfrak{H}}$. Thus, in this case $\mathfrak{F} = I_{\sigma}^{\tau} \text{ form } G$, where G is a monolithic $\bar{\tau}$ -minimal non- $\mathfrak{H}$ -group with monolith $P = G^{\mathfrak{H}}$, satisfying condition 2).

Let P be a σ -primary group. By lemma 28 the group G is a $\bar{\tau}$ -minimal non- $\mathfrak{H}$ -group. By lemma 23 the formation $\mathfrak{M}$ has an integrated definition m such that $m(\sigma_i) = \tau \text{ form}(\mathfrak{X} \cup \{K/Q\})$, where $\mathfrak{X}$ is the set of all proper $\bar{\tau}$ -subgroups of K , $m(\sigma_j) = \tau \text{ form}(G/O_{\sigma_j, \sigma_j}(G))$ for all $\sigma_j \in \sigma(G) \setminus \{\sigma_i\}$ and $m(\sigma_j) = \emptyset$ for all $\sigma_j \notin \sigma(G)$. By condition $\mathfrak{M} \subseteq \mathfrak{H}$. Hence, $m \leq h$. Since $m(\sigma_j) = \tau \text{ form}(G/O_{\sigma_j, \sigma_j}(G))$ for all $\sigma_j \in \sigma(G) \setminus \{\sigma_i\}$, it follows that for all such σ_j we have $G/O_{\sigma_j, \sigma_j}(G) \in h(\sigma_j)$. Let us show that $G/O_{\sigma_j, \sigma_j}(G) \simeq K/O_{\sigma_j, \sigma_j}(K)$. Indeed, since $\sigma_i \neq \sigma_j$, we have $O_{\sigma_j, \sigma_j}(G)/P = O_{\sigma_j, \sigma_j}(G/P)$. Consequently,

$$K/O_{\sigma_j, \sigma_j}(K) \simeq (G/P)/O_{\sigma_j, \sigma_j}(G/P) = (G/P)/(O_{\sigma_j, \sigma_j}(G)/P) \simeq G/O_{\sigma_j, \sigma_j}(G).$$

Therefore, for all $\sigma_j \in \sigma(G) \setminus \{\sigma_i\}$ we have $K/O_{\sigma_j, \sigma_j}(K) \in h(\sigma_j)$. Let us show that $P = G^{\mathfrak{H}}$. It is clear that $G \notin \mathfrak{H}$. If Q is a non- σ -primary group, then $O_{\sigma_j, \sigma_j}(K) = 1$. Consequently, $G/P \simeq K \in h(\sigma_j) \subseteq \mathfrak{H}$. If Q is a σ_j -group, then $O_{\sigma_j, \sigma_j}(K) = O_{\sigma_j}(K)$. Hence, $K/O_{\sigma_j, \sigma_j}(K) = K/O_{\sigma_j}(K) \in h(\sigma_j) = \mathfrak{G}_{\sigma_j} h(\sigma_j)$. Consequently,

$$G/P \simeq K \in \mathfrak{G}_{\sigma_j} h(\sigma_j) = h(\sigma_j) \subseteq \mathfrak{H}.$$

Thus, $P = G^{\mathfrak{H}}$ and G satisfies condition 3).

Consider now the case when $f(\sigma_i)$ is an Abelian formation. Applying lemma 27, we see that $\mathfrak{F} = I_{\sigma}^{\tau} \text{ form } G$, where G is a monolithic $\bar{\tau}$ -minimal non- $(\mathfrak{G}_{\sigma_i} \mathfrak{G}_{\sigma_i} f(\sigma_i))$ -group with monolith $P = G^{\mathfrak{H}}$, such that $\sigma_i \in \sigma(P)$, and either $P = G^{\mathfrak{G}_{\sigma_i} f(\sigma_i)}$ is a non- σ -primary group and G is a $\bar{\tau}$ -minimal non- $(\mathfrak{G}_{\sigma_i} f(\sigma_i))$ -group, or $G = P \rtimes H$, where $P = C_G(P)$ is a p -group, $p \in \sigma_i$, and H satisfies one of the following conditions: 1) H is a monolithic $\bar{\tau}$ -minimal non- $(\mathfrak{G}_{\sigma_i} f(\sigma_i))$ -group with monolith $Q = H^{\mathfrak{G}_{\sigma_i} f(\sigma_i)} \subsetneq \Phi(H)$, such that $\sigma_i \notin \sigma(Q)$; 2) H is a minimal non- $f(\sigma_i)$ -group of one of the following types: a) the quaternion group of order 8, if $2 \notin \sigma_i$; b) an extraspecial group of order q^3 of prime odd exponent $q \notin \sigma_i$; c) a cyclic q -group, $q \notin \sigma_i$.

Let P be a non- σ -primary group. By lemma 25 the formation $\mathfrak{M}$ has an integrated σ -local definition m such that $m(\sigma_j) = \tau \text{ form}(\mathfrak{X} \cup \{G/P\})$ for all $\sigma_j \in \sigma(P)$. By condition $\mathfrak{M} \subseteq \mathfrak{H}$. Consequently, $m \leq h$. Thus, for all $\sigma_j \in \sigma(P)$ it holds that all proper $\bar{\tau}$ -subgroups of G belong to $h(\sigma_j)$ and $G/P \in h(\sigma_j)$. Suppose that there exists $\sigma_j \in \sigma(P)$ such that $G \in h(\sigma_j)$. Then since the σ -local definition h is integrated, we have $G \in \mathfrak{H} \subseteq \mathfrak{G}_{\sigma_i} h(\sigma_i)$. The obtained contradiction shows that $P = G^{h(\sigma_j)}$ for all $\sigma_j \in \sigma(P)$. It is also clear that $P = G^{\mathfrak{H}}$. Thus, in this case $\mathfrak{F} = I_{\sigma}^{\tau} \text{ form } G$, where G is a monolithic $\bar{\tau}$ -minimal non- $\mathfrak{H}$ -group with monolith $P = G^{\mathfrak{H}}$, satisfying condition 2).

Let P be a σ_i -group. Then G is a $\bar{\tau}$ -minimal non- $\mathfrak{H}$ -group by lemma 28.

Let H satisfy condition 1). Then by lemma 23 the formation $\mathfrak{M}$ has an integrated definition m such that $m(\sigma_i) = \tau \text{ form}(\mathfrak{X} \cup \{H/Q\})$, where $\mathfrak{X}$ is the set of all proper $\bar{\tau}$ -subgroups of the group H . By condition $\mathfrak{M} \subseteq \mathfrak{H}$. Hence, $m \leq h$. Consequently, $\mathfrak{X} \cup \{H/Q\} \subseteq h(\sigma_i) = \mathfrak{G}_{\sigma_i} f(\sigma_i)$. If $H \in h(\sigma_i)$, then $G/P \simeq H \in h(\sigma_i)$. Hence,

$$G \in \mathfrak{G}_{\sigma_i} h(\sigma_i) = h(\sigma_i) \subseteq \mathfrak{H}.$$

The obtained contradiction shows that H is a $\bar{\tau}$ -minimal non- $h(\sigma_i)$ -group and $H/Q \in h(\sigma_i)$. Hence, $Q = H^{h(\sigma_i)}$ and H satisfies condition 3) of the theorem.

If H satisfies condition 2), then, obviously, H satisfies condition 3) of the theorem.

Let us show that $P = G^{\mathfrak{H}}$. Obviously, $G \notin \mathfrak{H}$. Suppose that $G/P \notin \mathfrak{H}$. Then $H \notin \mathfrak{H}$. Consequently, $\mathfrak{F} = I_G^{\tau}$ form $G = I_G^{\tau}$ form H . Applying now lemma 5, we see that

$$f(\sigma_i) = \tau \text{ form}(G / O_{\sigma_i, \sigma_i}(G)) = \tau \text{ form}(G/P) = \tau \text{ form } H = \tau \text{ form}(H / O_{\sigma_i, \sigma_i}(H)).$$

If the group H satisfies condition 1), then by lemma 8 the formation $\tau \text{ form}(\mathfrak{X} \cup \{H/Q\})$ is the unique maximal τ -closed subformation of $\tau \text{ form } H$, where $\mathfrak{X}$ is the set of all proper $\bar{\tau}$ -subgroups of H . Since Q is a σ_i -group, we have $Q \subseteq O_{\sigma_i, \sigma_i}(H)$. Hence, $H / O_{\sigma_i, \sigma_i}(H) \in \tau \text{ form}(\mathfrak{X} \cup \{H/Q\})$. Thus,

$$H / O_{\sigma_i, \sigma_i}(H) \in \tau \text{ form}(\mathfrak{X} \cup \{H/Q\}) \subset \tau \text{ form } H = \tau \text{ form}(H / O_{\sigma_i, \sigma_i}(H)).$$

Contradiction.

Now let H satisfy one of conditions 2). Then H is a σ -primary group, i. e. a σ_i -group for some $\sigma_i \in \sigma$. Consequently, $H = O_{\sigma_i}(H) = O_{\sigma_i, \sigma_i}(H)$. Hence, $f(\sigma_i) = \tau \text{ form } H = \tau \text{ form}(H / O_{\sigma_i, \sigma_i}(H)) = (1)$. The obtained contradiction shows that $G/P \in \mathfrak{H}$. Therefore, $P = G^{\mathfrak{H}}$.

Sufficiency. Let f be the smallest τ -valued σ -local definition of $\mathfrak{F}$. To prove sufficiency, by lemma 11 it suffices to establish that $f(\sigma_j)$ is a minimal τ -closed non- $h(\sigma_j)$ -formation for every $\sigma_j \in \sigma(P)$.

If the group G satisfies condition 1), then $\mathfrak{F}$ is a minimal τ -closed σ -local non- $\mathfrak{H}$ -formation by lemma 21.

Let G satisfy condition 2) and $\sigma_j \in \sigma(P)$. Then by lemma 5 we have

$$f(\sigma_j) = \tau \text{ form}(G / O_{\sigma_j, \sigma_j}(G)) = \tau \text{ form } G.$$

Let $\mathfrak{X}$ be the set of all proper $\bar{\tau}$ -subgroups of G . Then by condition $\mathfrak{X} \cup \{G/P\} \subseteq h(\sigma_j)$. But $\tau \text{ form}(\mathfrak{X} \cup \{G/P\})$ is the unique maximal τ -closed subformation of $\tau \text{ form } G$ by lemma 8. It is also clear that G does not belong to $h(\sigma_j)$. Hence, $f(\sigma_j)$ is a minimal τ -closed non- $h(\sigma_j)$ -formation.

Now let G satisfy condition 3) and $\Phi(K) = 1$. Then

$$f(\sigma_i) = \tau \text{ form}(G / O_{\sigma_i, \sigma_i}(G)) = \tau \text{ form } K.$$

Let $\mathfrak{X}$ be the set of all proper $\bar{\tau}$ -subgroups of K . Then by condition $\mathfrak{X} \subseteq h(\sigma_i)$ and $K/Q \in h(\sigma_i)$. Hence, $\tau \text{ form}(\mathfrak{X} \cup \{K/Q\}) \subseteq h(\sigma_i)$. By lemma 8 $\tau \text{ form}(\mathfrak{X} \cup \{K/Q\})$ is the unique maximal τ -closed subformation of $f(\sigma_i) = \tau \text{ form } K$. Hence, every proper τ -closed subformation of $f(\sigma_i)$ is contained in $h(\sigma_i)$. If $f(\sigma_i) \subseteq h(\sigma_i)$, then $G/P \in h(\sigma_i)$. Hence, by lemma 7 we have $G \in \mathfrak{H}$. Contradiction. So, $f(\sigma_i) \not\subseteq h(\sigma_i)$. Thus, $f(\sigma_i)$ is a minimal τ -closed non- $h(\sigma_i)$ -formation.

Let K satisfy condition a), i. e. K is the quaternion group of order 8. Then by lemma 17 $f(\sigma_i) = \tau \text{ form } K$ is a minimal non-Abelian formation. Let $\mathfrak{M}$ be a maximal subformation of $f(\sigma_i)$. It is easy to see that $\mathfrak{M}$ is the formation of Abelian groups of exponent dividing 4, and $\mathfrak{M}$ is generated by any Abelian group of exponent 4. Among the subgroups of K there is a proper Abelian normal subgroup of exponent 4. Hence, $\mathfrak{M} \subseteq \text{form } \mathfrak{X}$, where $\mathfrak{X}$ is the set of all proper subgroups of K ; moreover, the subgroups of K are Abelian and their exponents divide 4. Consequently, $\text{form } \mathfrak{X} \subseteq \mathfrak{M}$. Thus, $\mathfrak{M} = \text{form } \mathfrak{X}$. But by condition $\mathfrak{X} \subseteq h(\sigma_i)$. Hence, $\mathfrak{M} \subseteq h(\sigma_i)$. So, $f(\sigma_i)$ is a minimal τ -closed non- $h(\sigma_i)$ -formation.

Now let K satisfy condition b), i. e. K is an extraspecial group of order q^3 of prime odd exponent q . Then by lemma 17 $\tau \text{ form } K = f(\sigma_i)$ is a minimal τ -closed non-Abelian formation. Let M be a maximal subformation of $f(\sigma_i)$. Then M is an Abelian formation and every non-identity group from M has exponent q . Since K is a minimal non- $h(\sigma_i)$ -group, $h(\sigma_i)$ contains groups of order q . But then $M \subseteq h(\sigma_i)$. Consequently, $f(\sigma_i)$ is a minimal non- $h(\sigma_i)$ -formation.

Finally, let K satisfy condition c), i. e. K is a cyclic q -group. Let $|K| = q^n$ and M be a maximal subgroup of K . Then $M \in h(\sigma_i)$, but $K \notin h(\sigma_i)$. Hence, $f(\sigma_i) \not\subseteq h(\sigma_i)$ and $\text{form } M \subseteq h(\sigma_i)$. To complete the proof of the theorem it is necessary only to establish that every proper subformation of $f(\sigma_i)$ is contained in $\text{form } M$. Suppose that in $f(\sigma_i)$ there is a proper subformation $\mathfrak{M}$ such that $\mathfrak{M} \not\subseteq \text{form } M$, and let A be a group of minimal order from $\mathfrak{M} \setminus \text{form } M$. Then A is monolithic. Since the formation $f(\sigma_i)$ is Abelian, A is a cyclic group

of order q^m . According to lemma 19 the exponent q^m of A divides the exponent q^n of K . If $m = n$, then $A \simeq K$ and then $f(\sigma_i) \subseteq \text{form } A \subseteq \mathfrak{M} \subset f(\sigma_i)$. The obtained contradiction shows that $m \neq n$. But the order of M is q^{n-1} . Hence, either A is isomorphic to some proper subgroup of M , or $A \simeq M$. This means that $A \in \text{form } M$. Contradiction. Consequently, $f(\sigma_i)$ is a minimal non- $h(\sigma_i)$ -formation.

By lemma 16 in each of the considered cases $f(\sigma_i)$ is a hereditary formation, therefore $f(\sigma_i)$ is a minimal τ -closed non- $h(\sigma_i)$ -formation. Applying now lemma 11 we conclude that $\mathfrak{F}$ is an $\mathfrak{H}_\sigma^\tau$ -critical formation. The theorem is proved.

Recall that a formation $\mathfrak{F}$ is called 2-multiply σ -local [10], if $\mathfrak{F}$ has a σ -local definition all of whose non-empty values are σ -local. It is clear that every 2-multiply σ -local formation is a σ -local formation of classical type.

Corollary 1. *Let $\mathfrak{H}$ be a 2-multiply σ -local formation and h be its canonical σ -local definition. Then $\mathfrak{F}$ is a minimal τ -closed σ -local non- $\mathfrak{H}$ -formation if and only if $\mathfrak{F} = l_\sigma^\tau \text{form } G$, where G is a monolithic $\bar{\tau}$ -minimal non- $\mathfrak{H}$ -group with monolith $P = G^{\mathfrak{H}}$, and one of the following conditions holds:*

- 1) $G = P$ is a simple σ_i -group such that $\sigma_i \notin \sigma(\mathfrak{H})$ and $\tau(G) \subseteq \{1, G\}$;
- 2) P is a non- σ -primary group and G is a $\bar{\tau}$ -minimal non- $h(\sigma_i)$ -group with $P = G^{h(\sigma_i)}$ for all $\sigma_i \in \sigma(P)$;
- 3) $G = P \rtimes K$, where $P = C_G(P)$ is a p -group, $p \in \sigma_i$, and K is a monolithic $\bar{\tau}$ -minimal non- $h(\sigma_i)$ -group with monolith $Q = K^{h(\sigma_i)} \not\subseteq \Phi(K)$, where $\sigma_i \notin \sigma(Q)$.

Corollary 2. *Let $\mathfrak{H} = LF_\sigma(h)$, where h is an integrated σ -local definition all of whose non-empty values are Abelian. Then $\mathfrak{F}$ is a minimal τ -closed σ -local non- $\mathfrak{H}$ -formation if and only if $\mathfrak{F} = l_\sigma^\tau \text{form } G$, where G is a monolithic $\bar{\tau}$ -minimal non- $\mathfrak{H}$ -group with monolith $P = G^{\mathfrak{H}}$, and one of the following conditions holds:*

- 1) $G = P$ is a simple σ_i -group such that $\sigma_i \notin \sigma(\mathfrak{H})$ and $\tau(G) \subseteq \{1, G\}$;
- 2) P is a non- σ -primary group and G is a $\bar{\tau}$ -minimal non- $h(\sigma_i)$ -group with $P = G^{h(\sigma_i)}$ for all $\sigma_i \in \sigma(P)$;
- 3) $G = P \rtimes K$, where $P = C_G(P)$ is a p -group, $p \in \sigma_i$, and K is a minimal non- $h(\sigma_i)$ -group of one of the following types:
 - a) the quaternion group of order 8, if $2 \notin \sigma_i$;
 - b) an extraspecial group of order q^3 of prime odd exponent $q \notin \sigma_i$;
 - c) a cyclic q -group, $q \notin \sigma_i$.

In the case when τ is the trivial subgroup functor, from theorem 1 we obtain following corollary.

Corollary 3 [7, theorem A]. *Let $\mathfrak{H}$ be a σ -local formation of classical type and h be its canonical σ -local definition. Then $\mathfrak{F}$ is a minimal σ -local non- $\mathfrak{H}$ -formation if and only if $\mathfrak{F} = l_\sigma \text{form } G$, where G is a monolithic group with monolith $P = G^{\mathfrak{H}}$, and one of the following conditions holds:*

- 1) $G = P$ is a simple σ_i -group, $\sigma_i \notin \sigma(\mathfrak{H})$;
- 2) P is a non- σ -primary group and $P = G^{h(\sigma_i)}$ for all $\sigma_i \in \sigma(P)$;
- 3) $G = P \rtimes K$, where $P = C_G(P)$ is a p -group, $p \in \sigma_i$, and K is either a monolithic group with monolith $Q = K^{h(\sigma_i)} \not\subseteq \Phi(K)$, where $\sigma_i \notin \sigma(Q)$, or a minimal non- $h(\sigma_i)$ -group of one of the following types:
 - a) the quaternion group of order 8, if $2 \notin \sigma_i$;
 - b) an extraspecial group of order q^3 of prime odd exponent $q \notin \sigma_i$;
 - c) a cyclic q -group, $q \notin \sigma_i$.

In the classical case, when $\sigma = \sigma^1 = \{\{2\}, \{3\}, \{5\}, \dots\}$ from theorem 1 we obtain following corollary.

Corollary 4 [4, theorem 2.2.10]. *Let $\mathfrak{H}$ be a formation of classical type and h be its canonical screen. Then $\mathfrak{F}$ is a minimal τ -closed local non- $\mathfrak{H}$ -formation if and only if $\mathfrak{F} = \tau^l \text{form } G$, where G is a monolithic $\bar{\tau}$ -minimal non- $\mathfrak{H}$ -group with monolith $P = G^{\mathfrak{H}}$, and one of the following conditions holds:*

- 1) $G = P$ is a group of prime order $p \notin \pi(\mathfrak{H})$;
- 2) P is a non-Abelian group and G is a $\bar{\tau}$ -minimal non- $h(p)$ -group with $P = G^{h(p)}$ for all $p \in \pi(P)$;
- 3) $G = P \rtimes H$, where $P = C_G(P)$ is a p -group, and H is either a monolithic $\bar{\tau}$ -minimal non- $h(p)$ -group with monolith $Q = H^{h(p)} \not\subseteq \Phi(H)$, where p does not divide $|Q|$, or a minimal non- $h(p)$ -group of one of the following types:
 - a) the quaternion group of order 8;
 - b) a non-Abelian group of order q^3 of prime odd exponent q ;
 - c) a cyclic primary group.

If, in addition, τ is the trivial subgroup functor, then we obtain following corollary.

Corollary 5 [3, theorem]. *Let $\mathfrak{H} = LF(h)$ be a formation of classical type and h be its maximal integrated local screen. The formation $\mathfrak{F}$ is an $\mathfrak{H}_l$ -critical formation if and only if $\mathfrak{F} = l \text{ form } G$, where G is a monolithic group with monolith $P = G^{\mathfrak{H}}$, and one of the following conditions holds:*

- 1) $P = G$ is a group of prime order;
- 2) P is a non-Abelian group and $P = G^{h(p)}$ for all $p \in \pi(P)$;
- 3) $G = P \rtimes K$, where $P = C_G(P)$ is a p -group, and K is a monolithic group with monolith $Q = K^{h(p)}$ (p does not divide $|Q|$), such that either $\Phi(K) = 1$ and $K^{h(q)} \subseteq Q$ for all $q \in \pi(Q)$, or K is a minimal non- $h(p)$ -group of one of the following types:
 - a) a cyclic primary group;
 - b) the quaternion group of order 8;
 - c) a non-Abelian group of order q^3 of prime odd exponent q .

References

1. Shemetkov LA. Screens of step formations. In: Chernikov SN, editor. *Proceedings of the 6th All-Union symposium on group theory; 1978 September 19–21; Cherkasy, Ukraine*. Kyiv: Naukova dumka; 1980. p. 37–50. Russian.
2. Skiba AN. On critical formations. *Izvestiya Akademii nauk Belorusskoi SSR. Seriya fiziko-matematicheskikh nauk*. 1980;4: 27–33. Russian.
3. Skiba AN. On critical formations. In: Chernikov NS, editor. *Infinite groups and related algebraic structures*. Kyiv: Institute of Mathematics of the Academy of Sciences of Ukraine; 1993. p. 258–268. Russian.
4. Skiba AN. *Algebra of formations*. Minsk: Belaruskaja navuka; 1997. 240 p. Russian.
5. Shemetkov LA, Skiba AN. *Formations of algebraic systems*. Moscow: Nauka; 1989. 256 p. (Sovremennaya algebra). Russian.
6. Safonova IN. On minimal σ -local non- $\mathfrak{H}$ -formations of finite groups. *Problems of Physics, Mathematics and Technics*. 2020; 4:105–112. Russian. EDN: MCGNXX.
7. Safonova IN. On critical σ -local formations of finite groups. *Proceedings of the Institute of Mathematics of the National Academy of Sciences of Belarus*. 2023;31(2):63–80. Russian. EDN: NEPDVI.
8. Safonova IN, Skrudz VV. On $\mathfrak{H}_\sigma^c$ -critical formations of finite groups. *Problems of Physics, Mathematics and Technics*. 2025; 3:99–111. Russian. EDN: BFOPLO.
9. Skiba AN. On σ -subnormal and σ -permutable subgroups of finite groups. *Journal of Algebra*. 2015;436:1–16. DOI: 10.1016/j.jalgebra.2015.04.010.
10. Skiba AN. On one generalization of the local formations. *Problems of Physics, Mathematics and Technics*. 2018;1:79–82. Russian. EDN: YWUMMV.
11. Chi Z, Safonov VG, Skiba AN. On one application of the theory of n -multiply σ -local formations of finite groups. *Problems of Physics, Mathematics and Technics*. 2018;2:85–88. Russian. EDN: XUFGEH.
12. Zhang C, Skiba AN. On Σ_i^σ -closed classes of finite groups. *Ukrainian Mathematical Journal*. 2019;70(12):1966–1977. DOI: 10.1007/s11253-019-01619-6.
13. Chi Z, Safonov VG, Skiba AN. On n -multiply σ -local formations of finite groups. *Communications in Algebra*. 2019;47(3): 957–968. DOI: 10.1080/00927872.2018.1498875.
14. Safonova IN, Safonov VG. On some properties of the lattice of totally σ -local formations of finite groups. *Journal of the Belarusian State University. Mathematics and Informatics*. 2020;3:6–16. DOI: 10.33581/2520-6508-2020-3-6-16.
15. Safonova IN. Some properties of n -multiply σ -local formations of finite groups. *Asian-European Journal of Mathematics*. 2022;15(7):2250138. DOI: 10.1142/S1793557122501388.
16. Safonova IN. On the τ -closedness of n -multiply σ -local formation. *Advances in Group Theory and Applications*. 2024;18: 123–136. DOI: 10.32037/agta-2024-005.
17. Doerk K, Hawkes To. *Finite soluble groups*. Berlin: Walter de Gruyter; 1992. X, 891 p. (de Gruyter expositions in mathematics; volume 4). DOI: 10.1515/9783110870138.
18. Shemetkov LA. *Formations of finite groups*. Moscow: Nauka; 1978. 272 p. (Sovremennaya algebra). Russian.

Received 03.11.2025 / revised 09.01.2026 / accepted 24.02.2026.

АННОТАЦИИ ДЕПОНИРОВАННЫХ В БГУ РАБОТ INDICATIVE ABSTRACTS OF THE PAPERS DEPOSITED IN BSU

УДК 51(075.8)

Прокашева В. А. **Высшая математика** : электрон. учеб.-метод. комплекс для спец. 6-05-1036-04 «Международная логистика» / В. А. Прокашева, М. В. Мартон, О. А. Велько ; БГУ. Минск, 2025. 284 с. Библиогр.: с. 281–282. Режим доступа: <https://elib.bsu.by/handle/123456789/337957>. Загл. с экрана. Деп. в БГУ 02.12.2025, № 015002122025. Текст : электронный.

Электронный учебно-методический комплекс (ЭУМК) по учебной дисциплине «Высшая математика» предназначен для студентов специальности 6-05-1036-04 «Международная логистика». В ЭУМК содержатся лекционный материал, задания для практических занятий, контрольные работы, примерный тематический план, содержание учебного материала, вопросы для подготовки к экзамену, список литературы.

УДК 622.83:519.6(075.8)

Журавков М. А. **Механико-математические модели современной геомеханики** : электрон. учеб.-метод. комплекс для спец. 7-06-0533-06 «Механика и математическое моделирование» / М. А. Журавков, С. Н. Лопатин ; БГУ. Минск, 2025. 70 с. : ил. Библиогр.: с. 68–70. Режим доступа: <https://elib.bsu.by/handle/123456789/338880>. Загл. с экрана. Деп. в БГУ 15.12.2025, № 015815122025. Текст : электронный.

Электронный учебно-методический комплекс (ЭУМК) по учебной дисциплине «Механико-математические модели современной геомеханики» предназначен для студентов специальности 7-06-0533-06 «Механика и математическое моделирование». В ЭУМК рассматриваются вопросы математического и численного моделирования геомеханических процессов с точки зрения теорий, подходов и методов современной механики. Для успешного овладения дисциплиной «Механико-математические модели современной геомеханики» необходимы базовые знания по таким дисциплинам, как «Механика сплошной среды», «Математические методы механики деформируемого твердого тела и основы механики разрушения», «Численные методы механики сплошной среды», «Вычислительная механика», «Компьютерная механика».

УДК 517.968(075.8)

Громак Е. В. **Интегральные уравнения** : электрон. учеб.-метод. комплекс для спец. 1-31 03 02 «Механика и математическое моделирование» / Е. В. Громак, В. Г. Кротов ; БГУ. Минск, 2025. 37 с. : ил. Библиогр.: с. 36–37. Режим доступа: <https://elib.bsu.by/handle/123456789/339401>. Загл. с экрана. Деп. в БГУ 24.12.2025, № 017523122025. Текст : электронный.

Электронный учебно-методический комплекс (ЭУМК) по учебной дисциплине «Интегральные уравнения» предназначен для студентов специальности 1-31 03 02 «Механика и математическое моделирование». ЭУМК содержит тексты лекций, планы практических и семинарских занятий, перечень контрольных вопросов, списки рекомендованной литературы.

УДК 519.6(075.8)

Численные методы : электрон. учеб.-метод. комплекс для спец. 6-05-0533-13 «Механика и математическое моделирование» / О. А. Архипенко, А. И. Азаров, М. В. Игнатенко, Т. С. Якименко ; БГУ. Минск : БГУ, 2025. 91 с. Библиогр.: с. 90–91. Режим доступа: <https://elib.bsu.by/handle/123456789/339452>. Загл. с экрана. Деп. в БГУ 29.12.2025, № 017929122025. Текст : электронный.

Электронный учебно-методический комплекс (ЭУМК) по учебной дисциплине «Численные методы» разработан в соответствии с образовательным стандартом 1-й ступени высшего образования для специальности 6-05-0533-13 «Механика и математическое моделирование». ЭУМК предназначен для информационно-методического обеспечения преподавания дисциплины «Численные методы» для студентов данной специальности. В ЭУМК содержится конспект лекций, перечень лабораторных занятий с материалами для работы, задания по управляемой самостоятельной работе.

УДК 51(075.8)(076.1) + 51:378.016(075.8)

Ляцкая А. В. Методы классификации и решения математических задач : электрон. учеб.-метод. комплекс для спец. 6-05-0533-06 «Математика» / А. В. Ляцкая, Н. В. Бровка ; БГУ. Минск, 2025. 37 с. : ил. Библиогр.: с. 35–36. Режим доступа: <https://elib.bsu.by/handle/123456789/339455>. Загл. с экрана. Деп. в БГУ 29.12.2025, № 018029122025. Текст : электронный.

Электронный учебно-методический комплекс (ЭУМК) по учебной дисциплине «Методы классификации и решения математических задач» предназначен для студентов специальности 6-05-0533-06 «Математика». ЭУМК содержит тексты лекций, планы практических и семинарских занятий, перечень контрольных вопросов, списки рекомендованной литературы.

УДК 51(075.8)(076.1) + 51:378.016(075.8)

Методы программирования : электрон. учеб.-метод. комплекс для спец. 6-05-0533-07 «Математика и компьютерные науки» / А. И. Кравчук, А. С. Кравчук, Е. В. Кремень, Ю. А. Кремень ; БГУ. Минск, 2025. 471 с. : ил. Библиогр.: с. 469–470. Режим доступа: <https://elib.bsu.by/handle/123456789/340202>. Загл. с экрана. Деп. в БГУ 20.01.2026, № 001020012026. Текст : электронный.

Электронный учебно-методический комплекс (ЭУМК) по учебной дисциплине «Методы программирования» предназначен для студентов специальности 6-05-0533-07 «Математика и компьютерные науки» (профилизация «Веб-программирование и интернет-технологии», «Математическое и программное обеспечение мобильных устройств», «Искусственный интеллект и математическая экономика»). ЭУМК содержит тексты лекций, планы лабораторных занятий, перечень контрольных вопросов, списки рекомендованной литературы.

УДК 004.432.045С#(075.8)

Кравчук А. И. Язык программирования С#. Основы ASP.NET : электрон. учеб.-метод. комплекс для спец. 6-05-0533-07 «Математика и компьютерные науки» / А. И. Кравчук, Е. В. Кремень, Ю. А. Кремень ; БГУ. Минск, 2025. 335 с. : ил. Библиогр.: с. 332. Режим доступа: <https://elib.bsu.by/handle/123456789/340208>. Загл. с экрана. Деп. в БГУ 20.01.2026, № 001120012026. Текст : электронный.

Электронный учебно-методический комплекс (ЭУМК) по учебной дисциплине «Язык программирования С#. Основы ASP.NET» предназначен для студентов специальности 6-05-0533-07 «Математика и компьютерные науки» (профилизация «Веб-программирование и интернет-технологии»). ЭУМК содержит тексты лекций, планы лабораторных занятий, перечень контрольных вопросов, списки рекомендованной литературы.

УДК 004.42(075.8)

Аленский Н. А. Методы программирования : электрон. учеб.-метод. комплекс для спец.: 6-05-0533-08 «Компьютерная математика и системный анализ», 6-05-0533-13 «Механика и математическое моделирование» / Н. А. Аленский, Д. В. Филимонов ; БГУ. Минск, 2026. 310 с. Библиогр.: с. 309–310. Режим доступа: <https://elib.bsu.by/handle/123456789/341588>. Загл. с экрана. Деп. в БГУ 13.02.2026, № 002913022026. Текст : электронный.

Электронный учебно-методический комплекс (ЭУМК) по учебной дисциплине «Методы программирования» предназначен для студентов первого курса специальностей 6-05-0533-08 «Компьютерная математика и системный анализ» и 6-05-0533-13 «Механика и математическое моделирование». ЭУМК

включает тексты лекций, темы лабораторных работ, типы заданий и требования к их выполнению, тренировочные упражнения и тесты, вопросы к экзаменам и зачетам, содержание дисциплины и список рекомендуемой литературы.

УДК 004(075.8)

Аленский Н. А. **Практикум по информатике** : электрон. учеб.-метод. комплекс для спец. 6-05-0533-06 «Математика» / Н. А. Аленский ; БГУ. Минск, 2026. 89 с. Библиогр.: с. 86–89. Режим доступа: <https://elib.bsu.by/handle/123456789/341592>. Загл. с экрана. Деп. в БГУ 13.02.2026, № 003013022026. Текст : электронный.

Электронный учебно-методический комплекс (ЭУМК) по учебной дисциплине «Практикум по информатике» предназначен для студентов четвертого курса специальности 6-05-0533-06 «Математика». ЭУМК включает тексты лекций, темы лабораторных работ, типы заданий и требования к их выполнению, тренировочные упражнения и тесты, вопросы к зачету, содержание дисциплины и список рекомендуемой литературы.

УДК 519.712(075.8) + 004.422.63(075.8)

Соболевская Е. П. **Алгоритмы и структуры данных** : электрон. учеб.-метод. комплекс для спец.: 6-05-0533-09 «Прикладная математика», 6-05-0533-10 «Информатика», 6-05-0533-11 «Прикладная информатика», 6-05-0533-12 «Кибербезопасность» : в 3 ч. Ч. 3 / Е. П. Соболевская, В. М. Котов, А. А. Буславский ; БГУ. Минск, 2026. 224 с. : ил. Библиогр.: с. 223–224. Режим доступа: <https://elib.bsu.by/handle/123456789/342018>. Загл. с экрана. Деп. в БГУ 18.02.2026, № 003517022026. Текст : электронный.

Электронный учебно-методический комплекс (ЭУМК) по учебной дисциплине «Алгоритмы и структуры данных» (часть 3) предназначен для студентов специальностей 6-05-0533-09 «Прикладная математика», 6-05-0533-10 «Информатика», 6-05-0533-11 «Прикладная информатика», 6-05-0533-12 «Кибербезопасность». В ЭУМК содержится материал по разделу «Специальные структуры данных». Изложение соответствует программе учебной дисциплины.

УДК 517.98(075.8)

Функциональный анализ : электрон. учеб.-метод. комплекс для спец.: 6-05-0533-07 «Математика и компьютерные науки», 6-05-0533-06 «Математика» / А. Б. Антонец, С. Г. Штин, А. С. Бедрицкий, М. М. Логиновская ; БГУ. Минск, 2026. 213 с. : ил. Библиогр.: с. 212–213. Режим доступа: <https://elib.bsu.by/handle/123456789/342268>. Загл. с экрана. Деп. в БГУ 23.02.2026, № 004220022026. Текст : электронный.

Электронный учебно-методический комплекс (ЭУМК) по учебной дисциплине «Функциональный анализ» предназначен для студентов специальностей 6-05-0533-07 «Математика и компьютерные науки» и 6-05-0533-06 «Математика». В ЭУМК содержатся лекционный материал, практический материал, раздел контроля знаний и вспомогательный материал.

УДК 517.98(075.8)

Чев Е. С. **Функциональный анализ** : электрон. учеб.-метод. комплекс для спец. 6-05-0533-12 «Кибербезопасность» / Е. С. Чев ; БГУ. Минск, 2026. 249 с. Библиогр.: с. 248–249. Режим доступа: <https://elib.bsu.by/handle/123456789/343253>. Загл. с экрана. Деп. в БГУ 04.03.2026, № 004826022026. Текст : электронный.

Электронный учебно-методический комплекс (ЭУМК) по учебной дисциплине «Функциональный анализ» рассчитан на студентов специальности 6-05-0533-12 «Кибербезопасность», составлен в соответствии с образовательным стандартом 1-й ступени высшего образования и предназначен для информационно-методического обеспечения преподавания дисциплины «Функциональный анализ» для студентов данной специальности. В ЭУМК содержатся конспект лекций, перечень лабораторных занятий с материалами для работы, задания по управляемой самостоятельной работе. Изложение соответствует учебной программе.

УДК 004.434.045Scala(075.8)

Блинов И. Н. **Функциональное программирование и Scala-технологии** : электрон. учеб.-метод. комплекс для спец. 7-06-0533-04 «Математика и компьютерные науки» / И. Н. Блинов ; БГУ. Минск, 2026. 171 с. : ил. Библиогр.: с. 169. Режим доступа: <https://elib.bsu.by/handle/123456789/343523>. Загл. с экрана. Деп. в БГУ 10.03.2026, № 006310032026. Текст : электронный.

Электронный учебно-методический комплекс (ЭУМК) по учебной дисциплине «Функциональное программирование и Scala-технологии» предназначен для магистрантов специальности 7-06-0533-04 «Математика и компьютерные науки» (профилизация «Веб-программирование и интернет-технологии»). ЭУМК содержит тексты лекций, планы лабораторных занятий, перечень контрольных вопросов, списки рекомендованной литературы.

УДК 004.415.2(075.8)

Блинов И. Н. Анализ и проектирование информационных систем : электрон. учеб.-метод. комплекс для спец. 7-06-0533-04 «Математика и компьютерные науки» / И. Н. Блинов ; БГУ. Минск : БГУ, 2025. 96 с. Библиогр.: с. 93–94. Режим доступа: <https://elib.bsu.by/handle/123456789/343525>. Загл. с экрана. Деп. в БГУ 10.03.2026, № 006410032026. Текст : электронный.

Электронный учебно-методический комплекс (ЭУМК) по учебной дисциплине «Анализ и проектирование информационных систем» предназначен для магистрантов специальности 7-06-0533-04 «Математика и компьютерные науки». ЭУМК содержит тексты лекций, планы лабораторных занятий, перечень контрольных вопросов, списки рекомендованной литературы.

УДК 004.432.045:004.738.5JavaEE(075.8)

Блинов И. Н. Технологии JavaEE : электрон. учеб.-метод. комплекс для спец. 7-06-0533-04 «Математика и компьютерные науки» / И. Н. Блинов ; БГУ. Минск, 2025. 480 с. : ил. Библиогр.: с. 480. Режим доступа: <https://elib.bsu.by/handle/123456789/343527>. Загл. с экрана. Деп. в БГУ 10.03.2026, № 006510032026. Текст : электронный.

Электронный учебно-методический комплекс (ЭУМК) по учебной дисциплине «Технологии JavaEE» предназначен для магистрантов специальности 7-06-0533-04 «Математика и компьютерные науки». ЭУМК содержит тексты лекций, планы лабораторных занятий, перечень контрольных вопросов, списки рекомендованной литературы.

УДК 004(075.8)

Велько О. А. Основы информационных технологий : электрон. учеб.-метод. комплекс для спец. 6-05-0312-01 «Политология» / О. А. Велько, М. В. Мартон ; БГУ. Минск, 2026. 128 с. : ил. Библиогр.: с. 128. Режим доступа: <https://elib.bsu.by/handle/123456789/343580>. Загл. с экрана. Деп. в БГУ 11.03.2026, № 006611032026. Текст : электронный.

Электронный учебно-методический комплекс (ЭУМК) по учебной дисциплине «Основы информационных технологий» предназначен для студентов специальности 6-05-0312-01 «Политология». В ЭУМК содержатся лекционный материал, планы лабораторных занятий, примерные задания для управляемой самостоятельной работы студентов, тестовые задания, контрольные работы, вопросы для подготовки к зачету, список литературы.

УДК 34:004(075.8)

Моисеева Н. А. Информационные технологии в юридической деятельности : электрон. учеб.-метод. комплекс для спец. 6-05-0421-01 «Правоведение» / Н. А. Моисеева, О. А. Велько, О. М. Матейко ; БГУ. Минск, 2026. 128 с. : ил. Библиогр.: с. 128. Режим доступа: <https://elib.bsu.by/handle/123456789/343607>. Загл. с экрана. Деп. в БГУ 12.03.2026, № 006711032026. Текст : электронный.

Электронный учебно-методический комплекс (ЭУМК) по учебной дисциплине «Информационные технологии в юридической деятельности» предназначен для студентов специальности 6-05-0421-01 «Правоведение». В ЭУМК содержатся лекционный материал, планы лабораторных занятий, примерные задания для управляемой самостоятельной работы студентов, тестовые задания, контрольные работы, вопросы для подготовки к зачету, список литературы.

УДК 519.2(075.8)

Теория вероятностей и математическая статистика : электрон. учеб.-метод. комплекс для спец.: 6-05-0311-01 «Экономика», 6-05-0311-05 «Экономическая информатика», 6-05-0411-02 «Финансы и кредит», 6-05-0412-01 «Менеджмент», 6-05-1036-02 «Экономическая безопасность», 6-05-1036-03 «Государственный аудит», 7-07-1036-01 «Международная экономика и торговля (с углубленным изучением китайского языка)» / Е. И. Васенкова, М. В. Дубатовская, Т. В. Маковецкая, С. В. Рогозин ; БГУ.

Минск, 2026. 215 с. Библиогр.: с. 205–207. Режим доступа: <https://elib.bsu.by/handle/123456789/344513>. Загл. с экрана. Деп. в БГУ 30.03.2026, № 009527032026. Текст : электронный.

Электронный учебно-методический комплекс (ЭУМК) предназначен для студентов экономического факультета БГУ. Содержание ЭУМК предполагает изучение следующих вопросов: элементов комбинаторики, случайных событий и действий над ними, случайных величин (одномерных и многомерных), функций случайных величин, закона больших чисел, центральной предельной теоремы, выборочного метода в математической статистике, точечного и интервального оценивания параметров, проверки статистических гипотез, регрессионного и дисперсионного анализа.

СОДЕРЖАНИЕ

КОМПЬЮТЕРНЫЙ АНАЛИЗ ДАННЫХ

<i>Харин Ю. С., Абламейко С. В.</i> Современные тенденции в компьютерном анализе данных.....	6
<i>Савелов М. П.</i> Асимптотическая неотрицательная коррелированность ряда статистик, используемых при проверке качества генераторов случайных чисел.....	16
<i>Серов А. А.</i> Надежность двухуровневого подхода к тестированию в наборе тестов NIST SP 800-22.....	22

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

<i>Афанасьев В. И.</i> Предельная теорема о сходимости к локальному времени броуновской экскурсии.....	36
<i>Воробейчиков С. Э., Буркатовская Ю. Б.</i> Оценивание параметров пуассоновского процесса второго порядка, управляемого марковским процессом, с использованием процедуры обнаружения разладки.....	53
<i>Волошко В. А.</i> Локальная информационная геометрия для статистического анализа двоичных цепей Маркова высокого порядка.....	61
<i>Харин А. Ю.</i> Последовательная статистическая проверка гипотез для дискретных случайных данных с блочной структурой.....	75

ДИСКРЕТНАЯ МАТЕМАТИКА И МАТЕМАТИЧЕСКАЯ КИБЕРНЕТИКА

<i>Ду Сичжоу, Саражинский Д. С., Капский Д. В.</i> Включение доступности быстрых зарядных станций в модель назначений городских транспортных потоков на основе алгоритма Франк – Вульфа.....	81
<i>Орлович Ю. Л., Шутро Н. А.</i> О хорошо ve -укрытых и хорошо ev -укрытых графах.....	94

АНАЛИЗ ДАННЫХ

<i>Адамовский Е. Р., Богуш Р. П.</i> Система детектирования дыма и саботажа на базе одноплатного компьютера для мониторинга пожаров в режиме реального времени.....	108
<i>Андренко К. В., Крощенко А. А., Головкин В. А., Крапивин Ю. Б.</i> Анализ управленческих решений с использованием малых языковых моделей.....	121
<i>Дин Аоди, Йе Шипинг, Недзьведь А. М., Аносов В. С.</i> Система распознавания походки в реальном времени на основе легковесной двухмодальной оценки положения.....	130
<i>Краснопрошин В. В., Старовойтов А. А.</i> Мультимодельный подход с векторным поиском в оперативном управлении критичным ИТ-сервисом.....	140
<i>Цветков В. Ю., Рабцевич В. В., Ма Цзюнь, Зорько М. И.</i> Выделение объектов изображений для автономной локализации беспилотных летательных аппаратов и построения карты окружающего пространства.....	156

МАТЕМАТИЧЕСКАЯ ЛОГИКА, АЛГЕБРА И ТЕОРИЯ ЧИСЕЛ

<i>Скрудь В. В., Сафонова И. Н.</i> О проблеме Шеметкова описания критических формаций в классе τ -замкнутых σ -локальных формаций.....	163
Аннотации депонированных в БГУ работ.....	179

CONTENTS

COMPUTER DATA ANALYSIS

<i>Kharin Yu. S., Ablameyko S. V.</i> Current trends in computer data analysis.....	6
<i>Savelov M. P.</i> Asymptotic non-negative correlation of some statistics used in testing the quality of random number generators.....	16
<i>Serov A. A.</i> Reliability of two-level testing approach in the NIST SP 800-22 test suite.....	22

PROBABILITY THEORY AND MATHEMATICAL STATISTICS

<i>Afanasyev V. I.</i> Limit theorem on convergence to the local time of a Brownian excursion.....	36
<i>Vorobeychikov S. E., Burkatovskaya Yu. B.</i> Estimation of the parameters of a second order Markov modulated Poisson process using change-point detection procedure.....	53
<i>Voloshko V. A.</i> Local information geometry for statistical analysis of high-order binary Markov chains.....	61
<i>Kharin A. Yu.</i> Sequential statistical hypotheses testing for discrete random data with block structure.....	75

DISCRETE MATHEMATICS AND MATHEMATICAL CYBERNETICS

<i>Du Sizhuo, Sarazhinsky D. S., Kapski D. V.</i> Incorporating fast charging stations accessibility into an urban traffic assignment model based on the Frank – Wolfe algorithm.....	81
<i>Orlovich Y. L., Shutro M. A.</i> On well <i>ve</i> -covered and well <i>ev</i> -covered graphs.....	94

DATA SCIENCE

<i>Adamovskiy Y. R., Bohush R. P.</i> Smoke and camera tampering detection system based on single-board computer for real-time fire monitoring.....	108
<i>Andrenko K. V., Kroshchanka A. A., Golovko V. A., Krapivin Yu. B.</i> Managerial decision analysis using small language models.....	121
<i>Din Aodi, Ye Shiping, Nedzved A. M., Anosov V. S.</i> Real-time gait recognition system based on lightweight dual-modal pose estimation.....	130
<i>Krasnoproshin V. V., Starovoitov A. A.</i> Multi-model approach with vector search in operational management critical IT service.....	140
<i>Tsviatkou V. Yu., Rabtsevich V. V., Ma Jun, Zorko M. I.</i> Detection of image objects for autonomous localisation of unmanned aerial vehicles and map formation of the surrounding space	156

MATHEMATICAL LOGIC, ALGEBRA AND NUMBER THEORY

<i>Skrundz V. V., Safonova I. N.</i> On Shemetkov's problem of describing critical formations in the class of τ -closed σ -local formations.....	163
Indicative abstracts of the papers deposited in BSU.....	179

Журнал включен Высшей аттестационной комиссией Республики Беларусь в Перечень научных изданий для опубликования результатов диссертационных исследований по физико-математическим наукам (в области математики и информатики), техническим наукам (в области информатики).

Журнал включен в наукометрические базы данных Scopus, Mathematical Reviews, Ulrichsweb, Google Scholar, zbMath, Russian Science Citation Index, ПИИЦ, China National Knowledge Infrastructure.

**Журнал Белорусского
государственного университета.
Математика. Информатика.
№ 1. 2026**

Учредитель:
Белорусский государственный университет

Юридический адрес: пр. Независимости, 4,
220030, г. Минск, Республика Беларусь.

Почтовый адрес: пр. Независимости, 4,
220030, г. Минск, Республика Беларусь.

Тел. (017) 259-72-34, (017) 259-72-38.

E-mail: jmathinf@bsu.by

URL: <https://journals.bsu.by/index.php/mathematics>

«Журнал Белорусского государственного
университета. Математика. Информатика»

издается с января 1969 г.

До 2017 г. выходил под названием «Вестник БГУ.

Серия 1, Физика. Математика. Информатика»

(ISSN 1561-834X).

Редакторы *М. Д. Баранова, О. А. Семенец*

Технические редакторы *А. Р. Голік,*

М. А. Панкратова

Корректор *Л. А. Меркуль*

Подписано в печать 31.03.2026.

Тираж 25 экз. Заказ 0000.

**Journal
of the Belarusian State University.
Mathematics and Informatics.
No. 1. 2026**

Founder:
Belarusian State University

Registered address: 4 Niezaliezhnasci Ave.,
Minsk 220030, Republic of Belarus.

Correspondence address: 4 Niezaliezhnasci Ave.,
Minsk 220030, Republic of Belarus.

Tel. (017) 259-72-34, (017) 259-72-38.

E-mail: jmathinf@bsu.by

URL: <https://journals.bsu.by/index.php/mathematics>

«Journal of the Belarusian State University.
Mathematics and Informatics»

published since January, 1969.

Until 2017 named «Vestnik BGU.

Seriya 1, Fizika. Matematika. Informatika»

(ISSN 1561-834X).

Editors *M. D. Baranova, O. A. Semenets*

Technical editors *A. R. Golik,*

M. A. Pankratova

Proofreader *L. A. Merkul'*

Signed print 31.03.2026.

Edition 25 copies. Order number 0000.