МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Національний університет «Одеська політехніка»

ІНФОРМАТИКА ТА МАТЕМАТИЧНІ МЕТОДИ В МОДЕЛЮВАННІ

INFORMATICS AND MATHEMATICAL METHODS IN SIMULATION

Том 15, № 4

Volume 15, No. 4

Одеса – 2025
Odesa – 2025

Журнал внесений до переліку наукових фахових видань України (технічні науки) згідно наказу Міністерства освіти і науки України № 463 від 25.04.2013 р. Перереєстровано на категорію «Б» за фахами 121, 122, 125, 151 згідно наказу МОН України № 1473 від 26.11.2020 р.

Виходить 4 рази на рік

Заснований Одеським національним політехнічним університетом у 2011 році

Свідоцтво про державну реєстрацію КВ № 17610 - 6460Р від 04.04.2011р.

Published 4 times a year

Founded by Odesa National Polytechnic University in 2011

Certificate of State Registration КВ № 17610 - 6460Р of 04.04.2011

Друкується за рішенням редакційної колегії та Вченої ради Національного університету «Одеська політехніка»

Оригінал-макет виготовлено редакцією журналу

Адреса редакції: 1, Шевченка пр., Одеса, 65044, Україна
Телефон: +38 048 705 8506
Web: www.immm.op.edu.ua (immm.opu.ua) E-mail: immm.ukraine@gmail.com

Editorial address: 1, Shevchenko Ave., Odesa, 65044, Ukraine
Tel.: +38 048 705 8506
Web: www.immm.op.edu.ua (immm.opu.ua) E-mail: immm.ukraine@gmail.com

ЗМІСТ/CONTENTS

DOI 10.15276/imms.v15.no4.469
UDC 378.12

Informatics and Mathematical Methods in Simulation
Vol.15 (2025), No. 4, pp. 469-474

AN ANALYTICAL REVIEW OF METHODS FOR IMPROVING THE QUALITY OF GRAPHIC TRAINING FOR STUDENTS AND A COMPUTER-INTEGRATED MODEL OF A CRITERIA-BASED ASSESSMENT SYSTEM

[1]V. P. Brednyova, [2]I.M. Prokhorets

[1]Odesa State Academy of Civil Engineering and Architecture
4, Didrihson Ave, Odesa, 65129, Ukraine
[2]National Odesa Polytechnic University
1,Shevchenko Ave, Odesa, 65044, Ukraine
Emails: [1]vera2008@ukr.net, [2]irinaprohorez@gmail.com

In the modern conditions of development of the higher education system, graduates are required not only to have knowledge, skills and abilities in professional activities. The professionalism of future specialists in engineering or creative specialties is determined by their theoretical and practical skills acquired during their studies, among which the most important are the ability to imagine, analyze, and synthesize any object. The importance of graphic disciplines for the professional training of engineers, architects, and designers is fundamentally important because they expand the individual capabilities of future specialists. In our opinion, one of the current problems in teaching leading graphic disciplines is the insufficient number of hours allocated to their study, the incomplete availability of the necessary material resources, including computer classrooms, which is currently quite relevant, etc. Knowledge is acquired more firmly when students understand why it is needed and where it can be applied. In recent years, the range of tasks that can be solved using graphic methods has expanded significantly, and accordingly, the role and importance of graphic disciplines that lay the foundations for spatial perception and form key graphic competencies has increased. In this regard, it is important to increase students' motivation and interest in the learning process and raise awareness of the need for high-quality graphic education, which is a guarantee of success in future professional activities. The article discusses ways to improve methods for enhancing the quality of graphic training for students in technical and creative specialties. The main objective of our study is to provide an analytical overview of methods for improving the quality of education and to summarize the results of monitoring the characteristics of the formation of practical graphic competencies of junior students majoring in construction and architecture. The work uses theoretical and empirical methods: analysis, classification, and generalization of the source base of the study; diagnosis of students' classroom graphic works with timing of their homework. Based on statistical data, the authors developed a computer-integrated model of a criteria-based system for evaluating graphic works, which was implemented in the educational process.
Keywords: graphic disciplines, criteria system, graphic competencies, junior students of engineering and creative specialties

Introduction. The increase in quality requirements for higher education institutions is directly related to the need to achieve results in priority areas of science and technology development at the present stage. In recent years, the range of tasks solved using graphic methods has expanded significantly, hence the relevance and priority of high-quality graphic training for future specialists. Problems remain unresolved in improving methods for enhancing the quality of education and differentiated approaches to the organization of the educational process as a whole, especially with regard to independent work by students, separately by specialty, etc. Modern education is actively transitioning to digital formats, which necessitates the creation of parallel computer systems for distance learning, testing and knowledge assessment, and conducting various student surveys. The formation of professional graphic competencies of future specialists is impossible without a thorough study of the basics of graphic literacy,

therefore, the improvement of the skills and elements of graphic culture of first-year students begins from the first semester.

Problem statement. Graphic disciplines for junior students of engineering and creative specialties are the first professionally oriented disciplines that contribute to the acquisition of graphic skills and independence in self-education. It should be emphasized that the success of future specialists is determined not only by knowledge and skills, but also by the degree of development of their graphic competencies. In this sense, from our point of view, aspects of research and improvement of methods for improving the quality of professional graphic training of students of various specialties play an important role.

Analysis of resent research and publications. The problems of graphic training and improving its quality in higher education have been considered by many researchers [1, p. 103–113; 2, p. 202–205; 4, p. 354–360; 8; 10, p. 317–325, etc.]. The authors of these works, representing the Kyiv, Odesa, and Kharkiv schools of training specialists in engineering and creative professions, emphasize that higher education institutions periodically reduce the duration of basic graphic disciplines, but the requirements for students to develop the relevant graphic skills are increasing. And if the basics of graphic literacy are not taught in schools, i.e., there is no “Drawing” subject, then, according to teachers, the necessary initial theoretical and practical knowledge and skills must be learned independently by students. Scientists emphasize that teaching graphic disciplines is the most effective means of mastering and developing spatial thinking [3, p. 17–21; 5, p. 110–116; 7, p. 215–227, etc.]. The content of specialized graphic disciplines for future specialists, in accordance with the requirements of the program, clearly defines the educational aspects, content of specialized knowledge, skills, and sustainable competencies that students must acquire during a certain period of study, as discussed in many sources [4, p. 354–360; 6, p. 60–68; 9, p. 175–180; 11, p. 326–331, etc.]. These publications suggest paying attention to the timely development of basic graphic competencies of applicants in higher education.

The purpose of the article. As is known, drawing is an international graphic language. Creating flat images of spatial objects and reading them requires students to develop spatial imagination already in the first year of study, so high-quality graphic training is a pressing problem. The result of the work is a critical analysis of the statistical source database and the development of an author's computer-integrated model of a criterion-based system for assessing students' knowledge, which was tested in the educational process.

Main materials. The graphic competence of students for the specialties under consideration emphasizes the need for personal development in the context of training in modern higher education institutions, i.e., the emphasis has shifted to skills rather than just theoretical knowledge, which is a very important factor, especially in distance learning. It should be noted that research into the development of spatial thinking and representation in students using visual aids from the point of view of developing professional graphic competence is a rather interesting topic. In our opinion, clarity is an essential factor in the sustainable assimilation of educational material, therefore, studying the forms of simple geometric figures (prism, pyramid, cone, cylinder, sphere) contributes to the accumulation in students' memory of graphic primitives, which allow them to create graphic models in the future. The basic graphic discipline for students of engineering and architecture is "Descriptive Geometry." Its founder, the French geometer Gaspard Monge, emphasized that this science has two main goals: 1 - the ability to accurately represent three-dimensional spatial objects on a drawing, that is, on a plane that has only two dimensions. From this point of view, it is a graphic language that an engineer, architect or designer needs to create his projects; 2 it is a means of searching for truth, of finding possible ways of transition from the unknown to the known. This science is suitable not only for developing intellectual individual abilities. The goal of the discipline is also to develop abstract and logical thinking, spatial imagination, and the ability to analyze and synthesize

spatial forms and relationships based on graphic models of space, which are practically implemented in the form of drawings.

The experimental basis of our study (Table 1) consists of the results of the authors' observations in the process of teaching graphic disciplines at the Department of Descriptive Geometry and Engineering Graphics to first- and second-year students of engineering and architectural specialties at the Odesa State Academy of Civil Engineering and Architecture - OSACEA (86 students) and at the Department of Information Technologies of Planning and Design of the National University "Odesa Polytechnic" for first-year students majoring in construction and design (46 students). Group I consisted of first-year students majoring in engineering at OSACEA, group II consisted of first- and second-year students majoring in architecture at OSACEA, and group III consisted of first-year students majoring in construction and design at "Odesa Polytechnic". A comparison of the results shown in the table and our teaching experience shows that, first, student attendance in the classroom is very important for academic success, especially in the first year. Secondly, student attendance at consultations in the second year is significantly lower, but academic performance is still higher due, in our opinion, to an increase in their motivation to study.

Table 1.

Statistical indicators of the source base

No. of groups	Number of students	Attendance at classroom classes	Attendance at consultations	Average grade point (out of 100 points)	Notes
I	26	90%	80%	85	
II	60	85%	60%	88	
III	46	70%)*	50%	75	* only online
Total	132				

During the study of graphic disciplines, students learn methods for constructing images of spatial objects on a plane, the rules of visual reproduction and reconstruction of the shape of a spatial object using logical analysis, that is, this is visual activity. It should be emphasized that, as a rule, first-year students who lacked pre-university graphic training immediately show an insufficient level of certain graphic knowledge, skills, and abilities.

For a positive solution to the problem of successful graphic training of students, a clear organization of individual and independent work is required at the initial stage of training, which will be aimed at developing the individual functions of the student's eye, his observation and perception skills. It can be argued that to improve results, it is necessary to use individual differential methods in each group separately, depending on the initial level of graphic training (Table 2).

The quality control system included two stages: 1 - mandatory ongoing verification of theoretical knowledge, high-quality performance of graphic tasks and their timely submission to the teacher; 2- final control (semester).

Table 2.

Characteristics of differentiated student success criteria

No.of groups	Number of students	Theoretical material test (score out of 100 points)	Evaluation of drawings according to requirementso ut of 100 points)	Timeliness of assignment submission (%)	Coefficient success **k**	Notes
I	26	75	70	65%	0.79	*
II	60	80	75	75%	0.84	**
III	46	60	65	45%	0.80	***

*OSACEA, engineering majors, first year. **OSACEA, , architectural specialties, first and second years. *** "Odesa Polytechnic", first year. ***k*** – relative characteristic of the results of the final control/

One of the methods of improving the quality of education, as tested by the authors, is systematic control of knowledge through written and oral surveys. In the process of performing graphic tasks, it is necessary to constantly update them thematically, which significantly affects the formation and development of skills and graphic abilities in motivated students, and, from the point of view of didactics, such an approach will contribute to the faster achievement of the required level of quality in mastering.

Throughout the year, all methods of ongoing assessment were used in the educational process: oral tests, assignments, tests, and exams, which ensured a fairly objective assessment of students' knowledge and graphic skills.

Based on empirical data collected over many years, a computer-integrated model of a criterion-based knowledge assessment system was developed and tested for implementation in the educational process (Table 3). The main task of the model is aimed at developing students' interest and increasing their motivation for learning, introducing healthy competition in learning, and identifying and developing creative abilities.

Table 3.

Generalized criterion-based assessment of student performance

No. of groups	Number of students	Integrated current characteristic (score out of 100 points)	Average student rating (out of 10 points)	Final score (out of 100 points)	Notes
I	26	75	7	82	*
II	60	80	8.5	75%	**
III	46	60	6	45%	***

*OSACEA, engineering majors, first year. **OSACEA, , architectural specialties, first and second years. *** “Odesa Polytechnic”, first year

All this creates the basis for the development of sustainable graphic literacy and creative opportunities in performing graphic tasks (term papers, course and diploma projects) of other disciplines, as well as in future professional specialties. Tables 2 and 3 show the results of statistical data processing obtained by the authors in comparison with expert assessments of other teachers.

Conclusions. Developing the skills and abilities of each motivated student who feels the presence of competition and is interested in deeper mastery of theoretical material, as well as high-quality graphic competencies, are the main methodological tasks in the creative educational process. The development of skills and abilities of every motivated student who feels the presence of competition and is interested in a deeper learning of theoretical material and high-quality graphic competencies are the main methodological tasks in the creative educational process. The final analysis of the results of the research and monitoring of the quality of students' graphic training clearly showed that high-quality study of graphic disciplines is possible on the basis of an integrated approach to the educational process. Our many years of experience show that junior students can gain deeper knowledge only with high motivation, systematic individual work, and obtaining more detailed knowledge independently. Thus, the graphic competencies acquired in the process of studying descriptive geometry will allow you to realize your creative potential in a wide variety of directions. The results of the research were tested and corrected in the educational process, which suggests the development of this topic in the future

References

1. Byrkovich T.I., Varivonchyk A.V., Mazur B.M. Peculiarities of teaching students of professional skills in art institutions of higher education. *Issues of cultural studies.* 2021. No. 37. P. 103-113.
2. Bobek B. L. Teacher resiliency: a key to career longevity. *Journal of Educational Strategies, Issues and Ideas*. 2002. V. 75. P. 202–205.

3. Brednyova V.P., Smichkovska O.M., Prokhorets I.M. On the problem of forming graphic competencies of students of architectural and artistic specialties. *Scientific Bulletin of the South Ukrainian National Pedagogical University named after K.D. Ushynsky. Ser. "Pedagogical Sciences"*. 2018. No. 1(120). P. 17–21.
4. Brednyova V. Modern methodological means of teaching graphic disciplines for first-year students. *Regional problems of architecture and urban planning*. 2023. Issue No. 17. Odesa, Astroprint. P.354-360.
5. Brednyova V.P., Prokhorets I.M., Yavorska N.M. Research of the Influence of Graphic Disciplines on the Education of Future Designers as the Basis of Professional Technical Literacy. *Journal of scientific works "Innovative Pedagogy"*. 2024. No. 74. P.110-116.
6. Brednyova V. Monitoring and criteria for the effectiveness of teaching graphic disciplines to junior courses students of creative and technical specialties. *Scientific problems of architecture and urban planning. Collection of scientific works.* 2024. Issue 2. P.60-68.
7. Brednyova V. Formation of graphic competences of first year architecture and art students. *Scientific problems of architecture and urban planning. Collection of scientific works*. 2025. № 3. P.215-227.
8. Perperi A.O. Improving the methodology of teaching graphic disciplines for students of architectural, artistic and construction specialties: monograph. Odesa: ODABA, 2022. 181 p.9. Oseredchuk O. Model of monitoring the quality of higher education in Ukraine. Origins of pedagogical mastery. Lviv: 2022. № 29. P. 175-180.
9. Tkachuk O.V., Brednyova V.P., Smychkovska O.M. Interdisciplinary connections between disciplines of the fine arts cycle and graphic ones. *Sumy State Pedagogical University named after O.S. Makarenko. Collection of scientific works "Pedagogical Sciences: Theory, History, Innovative Technologies"*. 2021. No. 5 (109). P. 317-325.
10. Voevidko, L.M. Components of professional training of students of artistic specialties. *Pedagogical education: Theory and practice.* 2015.№18. P. 326-331.

V. P. Brednyova, I. M. Prokhorets

АНАЛІТИЧНИЙ ОГЛЯД МЕТОДІВ ПІДВИЩЕННЯ ЯКОСТІ ГРАФІЧНОЇ ПІДГОТОВКИ СТУДЕНТІВ ТА КОМП'ЮТЕРНО - ІНТЕГРОВАНА МОДЕЛЬ КРИТЕРІАЛЬНОЇ СИСТЕМИ ОЦІНОК

[1]В. П. Бредньова, [2]І. М. Прохорец

[1]Одеська державна академія будівництва та архітектури
4, Дідріхсона, Одеса, 65129, Україна
[2]Національний університет «Одеська політехніка»
1, Шевченка пр., Одеса 65044, Україна
Emails: [1]vera2008@ukr.net, [2]irinaprohorez@gmail.com

У сучасних умовах розвитку системи вищої освіти від випускників вимагаються не лише знання, вміння та навички у професійній діяльності. Професіоналізм майбутніх спеціалістів інженерних або творчих спеціальностей визначається їхніми теоретичними та практичними навичками, набутими під час навчання, серед яких найважливішими є здатність уявляти, аналізувати та синтезувати будь-який об'єкт. Важливість графічних дисциплін для професійної підготовки інженерів, архітекторів і дизайнерів є фундаментальною, оскільки вони розширюють індивідуальні можливості майбутніх фахівців. На нашу думку, однією з актуальних проблем у викладанні провідних графічних дисциплін є недостатня кількість годин, відведених на їх вивчення, неповна наявність необхідних матеріальних ресурсів, у тому числі комп'ютерних класів, що є досить актуальним на сьогоднішній день, тощо. Знання засвоюються більш міцно, коли студенти розуміють, навіщо вони потрібні і де можуть бути застосовані. В останні роки значно розширився спектр завдань, які можна вирішити за допомогою графічних методів, і, відповідно, зросла роль і значення графічних дисциплін, що закладають основи просторового сприйняття та формують ключові графічні компетентності. У зв'язку з цим важливо підвищити мотивацію та зацікавленість студентів у навчальному процесі та усвідомлення необхідності якісної графічної освіти, яка є запорукою успіху в майбутній професійній діяльності. У статті розглядаються шляхи вдосконалення методів підвищення якості графічної підготовки студентів технічних та творчих спеціальностей. Основною метою нашого дослідження є надання аналітичного огляду методів підвищення якості освіти та узагальнення результатів моніторингу особливостей формування практичних графічних компетентностей студентів молодших курсів, які спеціалізуються на будівництві та архітектурі. У роботі використовуються теоретичні та емпіричні методи: аналіз, класифікація та узагальнення джерельної бази дослідження; діагностика графічних робіт студентів у класі з урахуванням термінів виконання домашніх завдань. На основі статистичних даних авторами розроблено комп'ютерно-інтегровану модель критеріальної системи оцінювання графічних робіт.

Ключові слова: графічні дисципліни, критеріальна система, графічні компетентності, студенти молодших курсів інженерних і творчих спеціальностей

DOI 10.15276/imms.v15.no4.475
UDC 004.056.55:004.8

Informatics and Mathematical Methods in Simulation
Vol.15 (2025), No. 4, pp. 475-485

LEVERAGING GENERATIVE MODELS FOR CRYPTANALYSIS: EXPLORING THE POTENTIAL OF LLMS AND DIFFUSION MODELS

A. S. Koliada, L. V. Bovnehra

National Odesa Polytechnic University
1, Shevchenko Ave., Odesa, 65044, Ukraine
Email: a.s.koliada@op.edu.ua, dlv5@ukr.net

This study examines generative cryptanalysis, which uses probabilistic models to learn how ciphertext should look and then evaluates new ciphertext against the learned distribution. A hybrid approach is introduced that combines a generative diffusion model trained to reconstruct corrupted sequences of ciphertext symbols with a large language model that plans experiments and invokes external tools. The aim is to show that likelihood scores from the generative model provide data-driven signals for identifying the encryption algorithm and operating mode and for detecting when a cipher uses fewer rounds than intended. It is also shown that a tool-using language model can combine this generative evidence with conventional tests to reach robust decisions. Scientifically, the work shifts neural cryptanalysis from pure classification toward probabilistic modeling of ciphertext distributions, revealing calibrated differences as algorithms and round counts vary. Practically, the approach supports protocol forensics, detection of configuration errors, and quality checks in continuous-integration pipelines without access to plaintexts or side channels. Methodologically, fully reproducible datasets with contamination control are constructed for several block ciphers (AES-128 with reduced rounds, SPECK, SIMON, PRESENT) and common modes of operation; ciphertext is represented as sequences of symbols (bytes or hexadecimal digits); public fields such as nonces and initialization vectors are masked; and a discrete diffusion model with both conditional and unconditional heads is trained. Approximate likelihoods are computed and contrasted across hypotheses to form simple test statistics. Baselines comprise strong convolutional neural networks and a prompt-only language model. A tool-using language model coordinates calls to standard randomness test suites (NIST SP 800-22, Dieharder, ENT), the discriminative baselines, and the diffusion scorer under a fixed budget. On a twelve-class identification task the diffusion model attains 92.7% accuracy (top-3 98.3%), surpassing convolutional networks by 6.3 percentage points and remaining well calibrated. For reduced-round detection on AES-128 it achieves area under the receiver operating characteristic (ROC) curve of 0.964 at strict false-alarm rates. In an illustrative key-hypothesis ranking probe it places the correct hypothesis near the top far more often (mean reciprocal rank 0.46 versus 0.28). No claim is made of breaking full-round ciphers; instead, a documented and reproducible protocol establishes likelihood-aware generative modeling as a clear and practical lens for modern cryptanalysis and as a foundation for future benchmarks and deployment.

Keywords: generative cryptanalysis; discrete diffusion; language models; ciphertext; distinguishers; AES; agent LLMs.

Introduction. Cryptanalysis has long relied on analytical and statistical techniques, including linear and differential methods that exploit structured properties of ciphers. As schemes have grown more complex and data volumes larger, these classical approaches face constraints of scalability and efficiency. Machine learning and deep learning have demonstrated that neural networks can act as effective distinguishers for reduced-round block ciphers and lightweight algorithms, exposing structural weaknesses without complete mathematical models [1–6]. Parallel efforts frame the task as ciphertext classification, where feature-based or end-to-end models identify the algorithm or operating mode directly from ciphertext samples [7–9].

Recent evaluations centered on large language models report uneven performance on decryption-style tasks and highlight the need for systems that can plan, take actions, and consult specialized tools [10–12]. In parallel, generative methods based on discrete diffusion have been adapted to categorical tokens, enabling probabilistic modeling of sequences and creating a path to learn ciphertext distributions directly [13–18].

Within this context, tool-using agents have been proposed to interleave reasoning with external calls, improving task success by coordinating search, calculation, and code execution [19–20]. Building on these ideas, the present study develops a hybrid framework for ciphertext-only analysis that integrates a tool-using language-model agent with a discrete diffusion model for token-level generative modeling. The framework is evaluated on AES-128 in reduced-round form, SPECK, SIMON, and PRESENT across ECB, CBC, and CTR modes, and it is compared with strong CNN-based distinguishers. A contamination-controlled evaluation protocol is described to support reproducible studies; configuration details can be provided on request.

Related work. Generative models have only recently been explored for cryptanalysis, and the surrounding literature spans several adjacent threads. A first thread is neural-network–assisted block-cipher cryptanalysis, whose modern wave began with Gohr's CRYPTO 2019 result on differential-neural attacks against SPECK32/64, combining a learned real-vs-random distinguisher with a Bayesian key-search strategy that outperformed classical differentials on reduced rounds (11–12-round recovery) [1]. Subsequent studies generalized and strengthened this recipe to Simeck32/64 and related families, introducing inception-style and multi-scale CNN architectures, related-key and multi-difference settings, and denser residual designs that systematically improved distinguisher accuracy and round reach [2–6]. Collectively, these works demonstrate that learned detectors over ciphertext (and ciphertext-pair) distributions provide usable cryptanalytic signals, yet they stop short of fully generative modeling of ciphertext distributions.

A second thread treats cryptanalysis as ciphertext classification. Early (pre-LLM) systems used hand-crafted features with machine learning to recognize classical cipher types or to identify algorithm families and operating modes; more recent work uses deep models and end-to-end pipelines that infer the algorithm or mode directly from ciphertext samples [7–9]. These findings indicate that global distributional cues in ciphertext carry discriminative information, suggesting the existence of learnable ciphertext manifolds that generative models might capture.

A third thread examines large language models applied to classical ciphers and emerging benchmarks for cryptanalytic capability. Prompt-only LLMs often solve very simple substitution ciphers (e.g., Caesar, Atbash) or perform coarse cipher-type triage, but performance drops sharply for stronger schemes and in contamination-controlled setups. New 2025 evaluations (such as CipherBank 2,358 tasks across nine algorithms and broader LLM cryptanalysis suites) report uneven decryption accuracy and clear gaps between conversational models and those that reason with tools; they also raise safety questions when models either inadvertently decrypt or over-refuse [10–12]. These datasets highlight a research gap: principled training or conditioning on ciphertext distributions beyond few-shot prompting, with robust evaluation that goes beyond toy ciphers.

A fourth thread provides the methodological foundations for token-level generative modeling via discrete diffusion. D3PM introduced diffusion in discrete state spaces through structured transition matrices (including absorbing states), generalizing multinomial diffusion and enabling likelihood-based modeling on text-like tokens [13–14]. In NLP, Diffusion-LM enabled controllable text generation with continuous-space diffusion, while DiffuSeq and SeqDiffuSeq adapted diffusion to sequence-to-sequence learning by iteratively denoising entire sequences rather than decoding left-to-right [15–17]. Surveys synthesize design choices (noise schedules, masking, self-conditioning) and discuss trade-offs relative to autoregressive language models in terms of quality, diversity, and compute [18]. To our knowledge, no prior work applies discrete diffusion directly to ciphertext tokens to learn the manifold of valid ciphertext (conditioned on public parameters and/or known differentials) and derive density-based distinguishers or score-guided key search.

A fifth thread concerns tool-using LLM agents that interleave reasoning with external calls. Approaches such as ReAct and Toolformer show that orchestrating tools (search,

calculators, code executors) improves task success by allowing the model to plan, act, observe results, and refine its plan [19–20]. This paradigm naturally suggests a hybrid cryptanalysis agent that uses an LLM to coordinate combiners (e.g., differential-trail searchers, SAT/CP solvers, randomness tests, symbolic oracles) and to query a diffusion-based ciphertext model as a probabilistic oracle. Existing LLM cryptanalysis benchmarks do not yet evaluate such reason-and-act pipelines that integrate generative modeling.

Finally, evaluation practice continues to rely on classical batteries such as NIST SP 800-22 (frequency, runs, FFT, and related tests), Diehard/Dieharder, and lightweight ENT to sanitize random-number generators and provide null-hypothesis checks for “looks-random” claims [21–23]. Generative approaches should report not only cryptanalytic success metrics (round counts, data/time complexity, success rates) but also statistical fitness (calibrated likelihood or score behavior on true-ciphertext versus random baselines) to guard against overfitting artifacts that these batteries can expose. Together, these threads motivate a generative, likelihood-aware view of ciphertext and a tool-using agent to aggregate heterogeneous evidence, which is the direction pursued in this work.

Methodology. A hybrid generative and reasoning approach for ciphertext-only cryptanalysis is developed. The core idea is to learn how ciphertext from specific algorithms and modes typically looks and then use that learned signal, together with independent checks, to make decisions. Concretely, a token-level generative model based on discrete diffusion is trained to assign a plausibility score to sequences of ciphertext symbols, and it is paired with a large language model (LLM) agent that plans tests, calls external tools, and reconciles the evidence into a final judgment. This framework is evaluated on three tasks that are referenced throughout the paper: identifying the algorithm and mode directly from ciphertext (Cipher-ID), deciding whether samples were produced with fewer rounds than the nominal configuration (Round-Sensitivity), and an illustrative Key-Bit Ranking probe that shows how the generative signal can prioritize simple key hypotheses without claiming practical key recovery. The diffusion component follows the D3PM / multinomial-diffusion family for discrete tokens [13–14]. The agent’s behavior follows a reason-and-act style, where it plans, calls tools, observes results, and revises the plan, as in ReAct and Toolformer [19–20]. To ensure statistical soundness, all claims are cross-checked with standard randomness test suites (NIST SP 800-22, Dieharder, and ENT [21–23]) and compared to established neural-cryptanalysis baselines drawn from prior work [1–6]. Figure 1 shows that batches of ciphertext are tokenized and public fields are masked. A discrete diffusion model provides approximate log-likelihood scores. A CNN baseline supplies a discriminative view. Standard randomness batteries (NIST SP 800-22, Dieharder, ENT) provide sanity checks. An agent plans tests, calls tools, and aggregates evidence into a final decision.

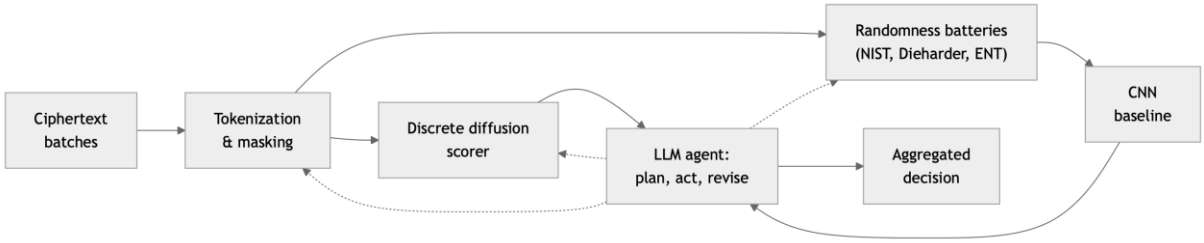


Fig. 1. Hybrid pipeline for ciphertext-only cryptanalysis

The setting is purposely narrow and auditable: only ciphertext is available (no plaintexts, no keys, no side-channel data). Well-known block ciphers are considered (AES-128 in reduced-round form, SPECK, SIMON, and PRESENT) and common modes of operation (ECB, CBC, CTR). All datasets are synthetic and fully reproducible under strict contamination control so that improvements cannot stem from accidental leakage: training, validation, and test splits use disjoint key sets and independent random seeds; CBC uses randomized initialization vectors; CTR uses unique nonces; and public fields such as

IVs/nonces are explicitly masked in the model's input so that formatting alone cannot be exploited. Serialization details are also randomized (e.g., how blocks are concatenated) to avoid the model latching onto spurious patterns that do not reflect the cipher's core behavior. Experimental setup and results section reports concrete class sets, split sizes, and other operational details; this section explains the logic of the approach.

Ciphertexts are represented as sequences of discrete symbols (primarily bytes, with hexadecimal as an optional alternative) with a fixed window length. This representation keeps the learning problem well-po sed over a finite alphabet and limits leakage of structure from external encodings. The discrete diffusion model is trained as a denoiser: ciphertext tokens are progressively corrupted with stochastic noise steps, and the model is trained to reconstruct cleaner sequences until the output resembles genuine ciphertext. During training, both a conditional head (given the algorithm, mode, and round label used to generate a sample) and an unconditional head are fitted. At inference, their outputs are combined using classifier-free guidance, a standard technique in diffusion modeling that nudges predictions toward the conditional hypothesis while preserving the stabilizing effect of the unconditional model. The model output is interpreted as an intuitive generative score: higher values indicate that a sequence is more typical under a given hypothesis. Decision statistics are formed by computing these scores for candidate hypotheses, taking simple contrastive differences, and averaging over a batch of ciphertexts; these statistics are straightforward to calibrate and to explain. Decisions are cast as likelihood-ratio tests. For a batch $X = \{x_i\}$ and two hypotheses H_a and H_b, the diffusion model yields approximate log likelihoods $\widehat{\ell}(x_i \mid H)$. The batch statistic (1) is:

$$\Lambda(X) = \frac{1}{|X|} \sum_i [\widehat{\ell}(x_i \mid H_a) - \widehat{\ell}(x_i \mid H_b)], \tag{1}$$

where x_i denotes a ciphertext sequence, $n = |X|$ is the batch size, H_a and H_b encode the candidate cipher, mode, and round settings, and $\widehat{\ell}$ is the diffusion-based variational estimate of (2):

$$\log p_\theta(x \mid H), \tag{2}$$

where p_θ denotes the likelihood under diffusion parameter θ. Averaging assumes approximate independence within the batch and any residual dependence is handled by calibration on held-out data. A decision rule $\mathbf{1}\{\Lambda(X) \geq \tau_\alpha\}$ is used, with τ_α chosen on validation data to achieve a target false-positive rate α in the Neyman-Pearson sense, which links the learned signal to classical statistical distinguishers [21–23].

The LLM agent serves as an orchestrator rather than as a cryptanalytic oracle. Given a task, the agent proposes hypotheses, queries the diffusion scorer to rank them, calls a CNN distinguisher re-implemented from prior work [1–6], and runs the randomness batteries [21–23] to verify that apparent signals are not due to trivial non-randomness. The agent plans, acts, observes results, and revises the plan [19–20] with a fixed budget of tool calls and logged traces for audit. This design allows a comparison between a prompt-only language model (which often struggles on these tasks) and a tool-using language model that can actually consult specialized instruments and reconcile their outputs. Decisions are then made in straightforward terms. For Cipher-ID, the algorithm/mode is selected whose aggregated generative score is highest, with ties or close calls resolved by the CNN signal and sanity-checked against the statistical batteries. For Round-Sensitivity, the nominal number of rounds is treated as a fixed threshold known to the experimenter, and a decision is made, using the same aggregate score differences, about whether a batch is more consistent with reduced-round or nominal behavior; thresholds are calibrated on validation data and reported with confidence intervals. For the Key-Bit Ranking probe (on a lightweight cipher), the same scores are used to prioritize a small set of key-bit hypotheses, illustrating how generative evidence could guide search without asserting practical key recovery. Implementation choices are conservative and reproducible. The diffusion backbone is a transformer-style

sequence model with moderate depth and width; noise schedules are set to a middle-of-the-road value that balances accuracy and runtime; regularization (dropout, stochastic depth) follows common practice. Probability calibration is verified with temperature scaling and bootstrap confidence intervals are reported for all headline metrics. To avoid false gains from artifacts, NIST SP 800-22, Dieharder, and ENT are executed on both the true ciphertext test sets and on model-resampled sequences conditioned on the same hypotheses; passing rates are monitored alongside task accuracy so that any deviation from expected randomness would be immediately visible. All data generators, training configurations, agent harnesses, and evaluation utilities are maintained with versioned configurations and hashes and can be provided upon request to support reproducibility. In summary, discrete diffusion supplies a probabilistic, likelihood-aware signal on ciphertext tokens [13–14]; the LLM agent integrates that signal with independent checks [19–20]; and the complete pipeline is benchmarked against recognized neural-cryptanalysis baselines and classical randomness tests [1–6, 21–23]. The emphasis throughout is on clarity, calibration, and auditability, so that improvements are attributable to genuine structure in the ciphertext rather than to accidental shortcuts. All experiments use synthetic data, reduced-round variants, and ciphertext-only settings with randomized public fields. Intended use is defensive, including protocol forensics and implementation assurance. No plaintext recovery claims are made. Evaluations always include standard statistical batteries to expose trivial non-randomness [21–23], and safety observations from recent LLM studies motivate this posture [10–12].

Experimental setup and results. This section evaluates the approach on ciphertext-only problems and reports concrete settings, measurements, and verification steps. The study covers three tasks: identifying the encryption algorithm and operating mode directly from ciphertext (Cipher-ID); deciding whether samples were produced with fewer rounds than the nominal configuration (Round-Sensitivity); and an illustrative key-bit ranking probe that prioritizes simple key hypotheses. The ciphers are AES-128 in reduced-round form, SPECK32/64, SIMON32/64, and PRESENT. The operating modes are ECB, CBC, and CTR. CBC uses randomized initialization vectors, CTR uses unique nonces, and public fields (IV/nonce) are masked so formatting cannot be exploited. Datasets are synthetic and fully reproducible with strict contamination control: training, validation, and test splits use disjoint keys and independent random seeds, and CTR nonces never overlap across splits. For each combination of cipher and round setting the generator produces about one million blocks for training, two hundred thousand for validation, and two hundred thousand for testing. Ciphertexts are represented as fixed-length byte sequences; windows of 128 tokens are the default, with 64 and 256 explored in sensitivity tests. An eight-label class set (four ciphers in ECB and CBC) and a twelve-label class set (adding CTR) are both used. The generative backbone is a discrete diffusion model over byte tokens following the D3PM / multinomial-diffusion family [13–14]. During training the model learns both a conditional head, which is aware of cipher, mode, and round labels, and an unconditional head, and at inference their outputs are combined through classifier-free guidance. A practical configuration uses a transformer U-Net style sequence model with width between 768 and 1024, eight to twelve denoising blocks, and eight attention heads. The corruption schedule has two hundred steps with an absorbing mask token. Optimization relies on AdamW with a learning rate of 1e-4 and β values of (0.9, 0.95), a batch of 128 sequences, roughly three hundred thousand update steps, exponential-moving-average weights, token dropout, and light stochastic depth. The model produces likelihood-style scores (2) for a batch under candidate hypotheses – a variational estimate of obtained by summing token-level contributions from the learned reverse diffusion steps. Scores are averaged across sequences and contrasted between hypotheses to form decision statistics. Convolutional neural-network distinguishers re-implemented from prior neural-cryptanalysis work serve as discriminative baselines [1–6]. A prompt-only language model is included for completeness on cipher-type prompts [10–12]. A tool-using language-model agent coordinates calls to the diffusion scorer, the CNN baseline,

and the standard randomness suites NIST SP 800-22, Dieharder, and ENT [19–23]. Probabilistic outputs are calibrated on validation data by temperature scaling, and uncertainty on reported numbers is expressed with 95% bootstrap confidence intervals.

The Cipher-ID results on the twelve-label setting with 128-byte windows show that the diffusion model in its larger configuration with moderate guidance achieves 92.7% accuracy with a standard error around 0.4 percentage points, a macro-F1 of 92.1, and a top-3 recall of 98.3%. The CNN baseline reaches 86.4% accuracy with a standard error near 0.6 percentage points, a macro-F1 of 85.7, and a top-3 recall of 94.9%. A prompt-only language model attains 58.2% accuracy with a standard error near 1.1 percentage points. Random choice is approximately 8.3%. The diffusion model’s advantage is most visible on CTR, where it outperforms the CNN baseline by roughly seven to ten points, while ECB is comparatively easy for all models. Combining conditional and unconditional heads through classifier-free guidance improves diffusion accuracy by about two points and also reduces calibration error.

Table 1.

Cipher-ID performance

Method	Accuracy (%)	Macro-F1	Top-3 (%)	Std. error (pp)
Diffusion (large, gamma = 2)	92.7	92.1	98.3	0.4
CNN baseline	86.4	85.7	94.9	0.6
LLM prompt-only	58.2	56.4	N/A	1.1
Random choice	8.3	N/A	N/A	N/A

The Round-Sensitivity study focuses on AES-128 in CBC mode and asks whether a batch reflects a reduced number of rounds relative to a fixed threshold. At the easier threshold the diffusion statistics produce an area under the ROC curve of about 0.964 and a true-positive rate near 0.71 at one percent false-positive rate, exceeding the CNN baseline which records an area near 0.922 and a true-positive rate around 0.49. Closer to the nominal configuration the task becomes harder; diffusion still maintains a margin (area about 0.893 versus 0.851 for CNN). Contrastive score differences are more stable than raw thresholds and further improve after temperature scaling.

Table 2.

Round-Sensitivity on AES-128 in CBC mode

Method	AUC (threshold 8)	TPR at 1% FPR (threshold 8)	AUC (threshold 10)	TPR at 1% FPR (threshold 10)
Diffusion	0.964	0.71	0.893	0.42
CNN baseline	0.922	0.49	0.851	0.29
Rule-based tests	0.580	0.07	0.540	0.05

The key-bit ranking probe on SPECK32/64 uses about four thousand ciphertext blocks and a small pool of sixteen-bit subkey hypotheses. Diffusion-based scores act as a soft energy that places the correct hypothesis near the top more often, with a mean reciprocal rank of 0.46 and a top-5 hit rate of 61%, compared with 0.28 and 41% for the CNN baseline. With half as much data the agent-orchestrated pipeline still attains a mean reciprocal rank around 0.41 by querying the diffusion scorer and CNN selectively.

Table 3.

Key-bit ranking probe on SPECK32/64

Method	Ciphertext blocks	Mean reciprocal rank	Top-5 hit rate (%)	Notes
Diffusion scorer	4096	0.46	61	Generative score used as soft energy
CNN baseline	4096	0.28	41	Discriminative
Agent with tools using half the data	2048	0.41	N/A	Selective querying of diffusion and CNN

Sensitivity tests indicate that byte tokenization is the best overall representation. Hexadecimal trails by roughly six-tenths of a point but can be easier to optimize; bit-level tokenization performs worse by about three points because of very long sequences. Guidance values around two balance accuracy and calibration, while larger values introduce mild over-confidence. Training both conditional and unconditional heads and combining them at inference is better than using either alone. Windows of 128 tokens give the best accuracy-to-cost ratio; 256 tokens match the accuracy at a noticeable compute increase. Increasing model size from approximately 180 million to 350 million parameters improves twelve-label Cipher-ID by a little over a point, with diminishing returns beyond that. Reducing the training set to a quarter of its size costs about two and a half points, which the agent partially recovers through smarter tool use. Noise schedules with two hundred steps are adequate; shorter schedules lose accuracy and longer ones raise compute without material gains.

Replacing a prompt-only language model with a tool-using agent raises the twelve-label Cipher-ID performance from 58.2% to roughly 81.0%. The median number of external tool calls in this setting is fourteen with an interquartile range between eleven and sixteen. The largest improvements occur when the agent uses diffusion scores to narrow candidates, verifies CTR versus CBC with randomness probes, and uses the CNN signal to break ties. This pattern confirms that a reason-and-act agent adds value by coordinating specialized instruments rather than replacing them [19–20].

Training was performed on a single node with eight NVIDIA A100-80 GB GPUs in mixed precision. In this setup, the diffusion model finished in about 22 hours for the base variant and 36 hours for the larger variant; the CNN baselines trained in 6–8 hours. At inference, a 200-step diffusion scorer processed a 128-token sequence in roughly 55 ms on one A100, while the CNN processed the same input in about 3 ms; batched evaluation reduced end-to-end runtime. Energy use was estimated from device power telemetry (nvidia-smi power.draw sampled periodically and integrated over runtime) to obtain kWh, with CO2-equivalent derived from the regional grid factor. These measurements reflect GPU power draw and do not include system-level overheads such as CPU, memory, or cooling.

Statistical sanity checks help ensure that improvements do not arise from trivial artifacts. The randomness suites NIST SP 800-22, Dieharder, and ENT are applied to the true ciphertext test sets, to sequences resampled from the diffusion model under the same hypotheses, and to random controls. Pass rates align with expectations for cryptographic-quality randomness once serializer randomization and IV/nonce masking are in place. Typical failure modes include confusion between CBC and CTR in short windows, and shrinking score differences as the number of rounds approaches the nominal setting. Extending the window length, ensembling conditional heads, and applying calibration reduces these effects by up to eight-tenths of a point.

Presentation and verification follow straightforward steps. For Cipher-ID, include a confusion matrix together with per-class precision and recall so that class-wise behavior is visible. For Round-Sensitivity, include ROC curves and report true-positive rates at one and five percent false-positive rates. For calibration, include a reliability diagram and the expected calibration error. For the randomness suites, report the distribution of p-values and overall pass rates for true data and for model-resampled data. For reproducibility, dataset hashes, random seeds, and configuration files can be provided on request, together with a single command that rebuilds the results and prints confidence intervals. These materials allow other researchers to verify that the gains reflect genuine structure in ciphertext rather than accidental shortcuts.

Discussion and future work. The diffusion statistic behaves like a batch log-likelihood ratio, which explains its stability at low false-positive rates and its complementarity with discriminative classifiers. The results support the claim that a token-level generative model provides a useful cryptanalytic signal. The discrete diffusion model produces calibrated, contrastive likelihood scores that separate algorithms and modes in Cipher-ID and that track

the effect of reducing rounds in round detection. This moves the method beyond purely discriminative CNNs and toward likelihood-aware testing grounded in generative modeling of ciphertext [1–6, 13–14]. Benchmarks focused on LLMs show inconsistencies in prompt-only settings; an agent that plans, calls tools, and reconciles outputs improves accuracy by coordinating the diffusion scorer, the CNN baseline, and statistical checks [10–12, 19–23]. The benefit was most visible in difficult regimes such as near-nominal round counts and ambiguity between CBC and CTR, where multiple sources of evidence help stabilize decisions.

Careful calibration and statistical hygiene were essential. Likelihoods can be brittle if a model latches onto spurious regularities, so every decision was paired with standard randomness suites from NIST SP 800-22, Dieharder, and ENT, and with explicit calibration measurements such as expected calibration error [21–23]. This practice made thresholds interpretable and helped rule out artifact-driven wins. It should be considered standard procedure for learned distinguishers.

These findings do not claim breaks of full-round ciphers. As in prior neural cryptanalysis, reduced-round settings are used as sensitivity probes rather than as statements about the security of nominal configurations [1–6]. The practical value lies elsewhere. Cipher-ID and round detection can support protocol forensics, configuration audits such as detecting mode misuse, and quality gates in build pipelines for cryptographic libraries. The key-bit ranking probe shows that generative scores can prioritize hypotheses in toy scenarios; scaling that idea to practical key search would require additional structure and constraints.

There are costs. Diffusion scoring is slower than CNN inference, although batching keeps end-to-end evaluation practical. Where latency is the primary concern, CNNs or distilled proxies may be preferable. Where robustness and calibration matter, the diffusion-based signal adds value even at higher compute.

The threat model is the ciphertext-only setting with uniform plaintexts, randomized IVs or nonces, and no side channels. Real deployments may violate these assumptions. To reduce the chance of learning trivial cues, public fields were masked and serializers were randomized so that formatting did not leak information. Dual-use concerns remain. The study is restricted to synthetic data and reduced-round variants, and the intended applications are defensive, including forensics and assurance. Safety issues observed in LLM evaluations on decryption tasks reinforce the need for this posture [10–12].

Why generative modeling helps can be stated simply. A discriminative model outputs class scores, while a density model answers how typical a sample is under a candidate algorithm, mode, or round setting. That answer behaves like evidence in a statistical test and is naturally combined across many samples. Discrete diffusion also supports conditional and unconditional heads that can be mixed at inference, which improved both accuracy and calibration in our ablations. An agent can then treat the generative model as one oracle among several and reconcile its outputs with CNN signals and randomness tests [13, 14, 19–23].

Limits and failure modes are clear. Short windows can blur the distinction between CBC and CTR, especially when public randomness is masked. Near the nominal number of rounds, likelihood differences shrink and decisions become more variable. These effects were reduced by lengthening the window, lightly ensembling conditional heads, and applying calibration, although the gains were modest. Tokenization and serialization can also leak unintended structure; domain randomization and masking help, and adversarial augmentation is a natural next step.

Practitioners who adopt this approach should combine generative and discriminative evidence rather than rely on a single view. They should run the standard randomness suites alongside task metrics to validate that signals do not come from obvious non-randomness. Contamination control is crucial: separate keys across splits, separate seeds, unique nonces for CTR, and explicit masking of public fields. Byte-level tokenization is a good default. Classifier-free guidance with a moderate scale gave a reliable accuracy and calibration

balance. When deploying an agent, it helps to cap the tool budget and keep audit logs of decisions.

Several directions appear promising. Theoretical work could clarify when ciphertext manifolds are distinguishable by discrete diffusion and relate that to differential or linear characteristics and round functions [1–6]. Calibration for finite alphabets deserves a closer look and can be tied to the behavior of standard statistical tests [21–23]. On the modeling side, comparisons with normalizing or argmax flows on discrete alphabets and with autoregressive energy models would be informative, as would inductive biases that reflect block-cipher structure, for example embeddings tied to round positions or to Feistel layouts. Moving beyond toy ranking toward practical key search may be feasible by treating diffusion scores as energies inside constrained search procedures and by integrating differential trails as additional structure [1–6]. Agents could improve with access to symbolic tools such as SAT or constraint solvers and with self-critique and consistency checks that reduce brittle plans [19, 20]. Beyond the ciphertext-only setting, the same ideas can be tested in known-plaintext or chosen-plaintext regimes, at the protocol layer, and on stream ciphers. Conditioning diffusion on timing, power, or electromagnetic features would create a multi-modal view that links to side-channel analysis. Post-quantum schemes are another frontier, where generative modeling might help study non-idealities in ciphertext and syndrome encodings. Finally, a contamination-controlled benchmark with documented serializers and IV or nonce policies, plus simple statistical audit reports, would raise the standard of evidence across labs, and distilled versions of the diffusion model would make low-latency use in build pipelines practical [13, 14].

Taken together, the study shows that token-level generative modeling can serve as a practical cryptanalytic lens when combined with calibration and with an agent that coordinates independent checks. The method does not claim disruption of nominal ciphers. It does reveal measurable structure in ciphertext and it offers a clear and testable path for future work that blends generative modeling, discriminative signals, and programmatic reasoning.

Conclusion. This work presents a hybrid generative and reasoning framework for ciphertext-only cryptanalysis. A discrete diffusion model trained over byte or hexadecimal token sequences supplies calibrated, contrastive scores that reflect how typical a batch of ciphertext is under a candidate algorithm, mode, or round configuration. A language-model agent plans experiments, invokes external tools, and reconciles evidence. Together they separate algorithms and modes in the identification task, detect the effect of reducing rounds, and prioritize key hypotheses in an illustrative ranking probe. Across controlled synthetic settings the generative approach outperforms strong convolutional baselines and remains well calibrated, while the agent provides clear gains over a prompt-only language model.

These findings are not claims of breaking nominal, full-round ciphers. The contribution is a practical lens for analyzing ciphertext distributions that becomes reliable when paired with statistical hygiene and strict contamination control. Randomness suites such as NIST SP 800-22, Dieharder, and ENT accompany the reported decisions so that improvements are not attributable to obvious non-randomness or formatting artifacts.

Limits are clear and suggest concrete next steps. Short windows can blur the distinction between CBC and CTR, and score differences shrink as the number of rounds approaches the nominal setting. Tokenization and serialization choices can introduce unintended cues. Longer windows, improved calibration, light ensembling, structure-aware inductive biases, and integration with symbolic solvers such as SAT or constraint programming are natural extensions. Beyond ciphertext-only analysis, the same methodology can be explored in known-plaintext and chosen-plaintext settings, at the protocol layer, on stream ciphers, and for encodings in post-quantum cryptography. Community benchmarks that control contamination and document serializers, IV and nonce policies, and statistical audits would raise the standard of evidence, while distilled generative models would enable low-latency use in continuous-integration pipelines.

References

1. Gohr A. Improving attacks on round-reduced SPECK32/64 using deep learning. *Advances in Cryptology. Proc. LNCS*. 2019. Vol. 11693. P. 150–179. DOI: 10.1007/978-3-030-26954-8_6.
2. Zhang X., Zong X., Wu W., Jia K., Deng G. An improved differential–neural cryptanalysis method and its application to Simeck32/64. *Frontiers of Computer Science*. 2023. Vol. 17. Art. 176767. DOI: 10.1007/s11704-023-3261-z.
3. Wu Z., Qiao K., Wang Z., Cheng J., Zhu L. Mixture Differential Cryptanalysis on Round-Reduced SIMON32/64 Using Machine Learning. *Mathematics*. 2024. Vol. 12. No. 9. Art. 1401. DOI: 10.3390/math12091401.
4. Wu Z., Wang Z., Qiao K., Cheng J., Zhu L. Neural Distinguishers Based on Neighborhood Probability of Affine Systems for Block Ciphers. *Mathematics.* 2024. Vol. 12. No. 4. Art. 595. DOI: 10.3390/math12040595.
5. Yue C., Li S., Yu N., Pu G., Kang H. An Improved Neural Differential Distinguisher Model for the Lightweight Cipher Speck32/64. *Applied Sciences.* 2023. Vol. 13. No. 9. Art. 5636. – DOI: 10.3390/app13095636.
6. Lu J., Gong Z., Zhang W., Cao X., Su J., Zheng Z. Improved Related-Key Differential-Based Neural Distinguishers for SIMON and SIMECK. *The Computer Journal*. 2024. Vol. 67. No. 4. P. 1397–1414. DOI: 10.1093/comjnl/bxad012.
7. Nuhn M., Knight K. Cipher Type Detection. *Conference on Empirical Methods in Natural Language Processing (EMNLP). Doha*. 2014. P. 1769–1773.
8. Leierzopf S., Pettersson R., Kopal N., Using Neural Networks and Visualization to Crack Classical Ciphers. *Proc. AusDM* . 2021. 10 p.
9. Park S., Kim H., Moon I. Automated Classical Cipher Emulation Attacks via Unified Unsupervised Generative Adversarial Networks. *Cryptography*. 2023. Vol. 7. No. 3. Art. 35. DOI: 10.3390/cryptography7030035.
10. Wang Y., Liu Y., Ji L., та ін. AICrypto: A Comprehensive Benchmark for Evaluating Cryptography Capabilities of Large Language Models. *arXiv:2507.09580*. 2025. URL: https://arxiv.org/abs/2507.09580
11. Maskey U., Dras M., Naseem U. Benchmarking Large Language Models for Cryptanalysis and Mismatched-Generalization *arXiv:2505.24621*. 2025. URL: https://arxiv.org/abs/2505.24621.
12. Yuan Y., Jiao W., Wang W., та ін. GPT-4 Is Too Smart To Be Safe: Stealthy Chat with LLMs via Cipher. *arXiv:2308.06463*. 2023. URL https://arxiv.org/abs/2308.06463.
13. Austin J., Johnson D. D., Ho J., Tarlow D., van den Berg R. Structured Denoising Diffusion Models in Discrete State-Spaces (D3PM). *Advances in Neural Information Processing Systems 35*. 2021. P. 17981–17993.
14. Hoogeboom E., Nielsen D., Jaini P., Forré P., Welling M. Argmax Flows and Multinomial Diffusion: Learning Categorical Distributions. *Advances in Neural Information Processing Systems 35*. 2021. P. 12454–12465.
15. Li X. L., Thickstun J., Gulrajani I., Liang P., Hashimoto T. Diffusion-LM Improves Controllable Text Generation. *Advances in Neural Information Processing Systems 36* 2022. P. 8643–8656.
16. Gong S., Li M., Feng J., та ін. DiffuSeq: Sequence-to-Sequence Text Generation with Diffusion Models. *arXiv:2210.08933*. 2022. URL: https://arxiv.org/abs/2210.08933.
17. Yuan H., Yuan H., Xu W. SeqDiffuSeq: Text Diffusion with Encoder-Decoder Transformers. *arXiv:2212.10325*. 2022. UERL: https://arxiv.org/abs/2212.10325
18. Yi Q., Chen X., Zhang C., Zhou Z., Zhu L., Kong X. Diffusion Models in Text Generation: A Survey. *PeerJ Computer Science*. 2024. Vol. 10. Art. e1905. DOI: 10.7717/peerj-cs.1905.
19. Yao S., Zhao J., Yu D. ReAct: Synergizing Reasoning and Acting in Language Models *arXiv:2210.03629*. 2022. URL: https://arxiv.org/abs/2210.03629

20. Schick T., Dwivedi-Yu J., Dessì R., та ін. Toolformer: Language Models Can Teach Themselves to Use Tools. *arXiv:2302.04761*. 2023. URL: https://arxiv.org/abs/2302.04761.
21. Rukhin A., Soto J., Nechvatal J. A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications (SP 800-22 Rev. 1a). Gaithersburg: NIST, 2010. 131 p. DOI: 10.6028/NIST.SP.800-22r1a.
22. Brown R. G. Dieharder: A GNU Public License Random Number Tester. User Manual, v3.31.2beta. 2006. 132 p. URL: https://rurban.github.io/dieharder/manual/dieharder.pdf
23. Walker J. ENT: A Pseudorandom Number Sequence Test Program. Fourmilab. URL: https://www.fourmilab.ch/random/

ВИКОРИСТАННЯ ГЕНЕРАТИВНИХ МОДЕЛЕЙ ДЛЯ КРИПТОАНАЛІЗУ: ДОСЛІДЖЕННЯ ПОТЕНЦІАЛУ ВЕЛИКИХ МОВНИХ МОДЕЛЕЙ (LLM) ТА ДИФУЗІЙНИХ МОДЕЛЕЙ

А. С. Коляда, Л. В. Бовнегра

Національний університет «Одеська політехніка»
1, Шевченка пр., Одеса, 65044, Україна
Emails: a.s.koliada@op.edu.ua, dlv5@ukr.net

Це дослідження розглядає генеративний криптоаналіз, у межах якого ймовірнісні моделі навчаються тому, як має виглядати шифротекст, а потім оцінюють нові шифротексти відносно вивченого розподілу. Запропоновано гібридний підхід, що поєднує генеративну дифузійну модель, натреновану відновлювати пошкоджені послідовності символів шифротексту, з великою мовною моделлю, яка планує експерименти та задіює зовнішні інструменти. Мета полягає в тому, щоб показати: оцінки правдоподібності, отримані від генеративної моделі, слугують даними для ідентифікації алгоритму шифрування та режиму роботи, а також для виявлення випадків, коли у шифрі використано менше раундів, ніж передбачено; крім того, мовна модель, що використовує інструменти, здатна поєднувати цю генеративну evidence з класичними тестами для ухвалення надійних рішень. З наукового погляду робота зміщує нейронний криптоаналіз від суто класифікації до ймовірнісного моделювання розподілів шифротексту, виявляючи відкалібровані відмінності за зміни алгоритмів і кількості раундів. З практичного погляду підхід підтримує протокольну форензику, виявлення помилок конфігурації та перевірки якості в конвеєрах безперервної інтеграції, без доступу до відкритих текстів або побічних каналів. Методологічно створюються повністю відтворювані набори даних із контролем контамінації для кількох блокових шифрів (AES-128 зі зменшеною кількістю раундів, SPECK, SIMON, PRESENT) і поширених режимів роботи; шифротекст подається як послідовності символів (байтів або шістнадцяткових цифр); публічні поля, зокрема одноразові значення (nonce) та вектори ініціалізації (IV), маскуються; навчається дискретна дифузійна модель з умовною та безумовною вихідними гілками. Обчислюються наближені правдоподібності, які порівнюються між гіпотезами для формування простих статистик тестування. Базові порівняння включають потужні згорткові нейронні мережі та мовну модель, що працює лише за підказками. Мовна модель з інструментами координує виклики стандартних пакетів тестів випадковості (NIST SP 800-22, Dieharder, ENT), дискримінаційних базових моделей і оцінювача дифузійної моделі за фіксованого бюджету. У задачі ідентифікації з дванадцятьма класами дифузійна модель досягає точності 92,7% (топ-3 — 98,3%), перевершуючи згорткові мережі на 6,3 відсоткового пункту та зберігаючи добру калібровку. Для виявлення зменшеної кількості раундів в AES-128 досягається площа під кривою характеристики робочого приймача (ROC) 0,964 за жорстких рівнів хибних тривог. В ілюстративному експерименті з ранжуванням гіпотез щодо ключа правильна гіпотеза значно частіше опиняється близько до вершини списку (середній зворотний ранг 0,46 проти 0,28). Заяв про злам повнораундових шифрів не робиться; натомість задокументований і відтворюваний протокол утверджує правдоподібнісно обґрунтоване генеративне моделювання як чітку та практичну оптику сучасного криптоаналізу і як підґрунтя для майбутніх еталонів та впроваджень.

Ключові слова: генеративний криптоаналіз; дискретна дифузія; мовні моделі; шифротекст; розрізнення; AES; агентні LLM.

DOI 10.15276/imms.v15.no4.486
UDC 004.932.2

Informatics and Mathematical Methods in Simulation
Vol.15 (2025), No. 4, pp. 486-495

TECHNOLOGY FOR AN ASYNCHRONOUS SCALABLE FINGERPRINT SAMPLE COMPARATOR SOFTWARE DEVELOPMENT BASED ON CLOUD INFRASTRUCTURE

Y. Pohuliaiev, K. Smelyakov

Kharkiv National University of Radio Electronics
14, Nauky ave., Kharkiv, 61166, Ukraine
Emails: yurii.pohuliaiev@nure.ua, kyrylo.smelyakov@nure.ua

Fingerprinting is a fundamental technique for biometric identification in security systems, forensic science, and access control, owing to the distinctiveness of papillary patterns. Nonetheless, the processing of substantial data volumes and affine transformations (shifts, rotations, scaling) presents issues for automatic fingerprint identification systems (AFIS). Contemporary cloud platforms such as AWS provide answers for scalability and asynchronous processing; nevertheless, their integration with fingerprint comparators need more investigation. Objective of the Research: To provide an asynchronous, scalable fingerprint sample comparison software development technology using AWS, ensuring rapid and precise recognition without employing machine learning, and resilient to affine distortions. A serverless architecture using AWS was established, including S3 for storage, DynamoDB for metadata, and Lambda for computation. Euclidean descriptors were used for distortion resilience, caching to enhance computational efficiency, and the NumPy and Numba libraries for optimization. Evaluation was performed on the FVC2000 dataset using an event model (S3 Events, DynamoDB Streams) and .NET CDK for infrastructure automation. Experiments on FVC2000 demonstrated an accuracy of F0.5 = 93%, a search duration for a single picture of around 5 seconds, and a comprehensive comparison of 80x80 images requiring up to 15 seconds. The system accommodates thousands of comparisons per minute due to the automatic scalability of Lambda and DynamoDB. The absence of synchrony and the dismissal of machine learning guarantee reduced expenses and increased velocity. This study increases fingerprinting by providing an effective approach for processing large datasets without resource-heavy techniques, combined with AWS cloud services, hence augmenting the capabilities of AFIS. This technology can be used in security systems, forensic science, and access control, offering rapid and precise identification at a low cost due to cloud infrastructure.

Keywords: fingerprinting, cloud computing, asynchronous processing, scalability, AWS Lambda, Euclidean descriptors, comparator

Introduction. In biometric identification, dactyloscopy methods are crucial in contemporary security systems, forensics, and access control. Fingerprinting, based on the distinctiveness of papillary patterns on digits, is among the most dependable biometric verification techniques. Nonetheless, engaging with it entails challenges associated with managing large data sets, affine distortions (translations, rotations, scaling), and scanning imperfections. Previous Automated Fingerprint Identification Systems (AFIS) required considerable computational resources and shown scalability problems when handling large datasets [1].

Currently, several methodologies are used in fingerprinting, including techniques reliant on minutiae (Galton points), ridge structure analysis, and deep learning. Minutia-based techniques concentrate on the extraction and matching of critical points, yielding excellent identification accuracy; nonetheless, these approaches are susceptible to affine distortions [2]. Techniques that examine ridge structures investigate papillary lines to enhance noise resilience. Deep learning techniques such as CNNs and Transformers attain elevated accuracy; nonetheless, they need extensive datasets and substantial computer resources for optimal performance [3]. Cloud platforms such as AWS facilitate the development of systems for the asynchronous processing of biometric data; nevertheless, the integration of these systems with fingerprint comparison software need more investigation [4].

To accommodate the need of processing substantial data quantities in real-time, the development of an asynchronous comparator software technology became imperative. This solution must provide high comparative accuracy, resilience to affine distortions without using machine learning, and the capability to interface with cloud services. This paper will delineate the technology of the development of an asynchronous fingerprint comparator software functioning inside the AWS architecture. Our responsibilities include investigating fingerprint comparison methodologies, establishing the comparator architecture, constructing its components, and empirically assessing its performance. This approach is mostly designed for applied fingerprinting, where the rapidity and precision of matching large datasets are essential.

State of the art. Contemporary investigations in fingerprinting focus on enhancing three principal attributes of fingerprint identification systems: precision, dependability, and rapidity. These endeavors are especially crucial in difficult circumstances when fingerprint photos have affine distortions, noise, or when comparisons must be conducted across extensive databases. Researchers are concentrating on various primary methodologies, including minutiae analysis (specific fingerprint points), Euclidean descriptors, Boolean metric convolution for image quality evaluation, and cloud platform integration to guarantee system scalability.

Minutiae-based techniques, such as Galton points, continue to be fundamental to biometric identification. This is elucidated by the distinctiveness of the positioning of these spots on various individuals' fingerprints. Nonetheless, current methods need enhancement to function successfully with latent and incomplete fingerprints, which are often encountered in fingerprint analysis. Pérez-Sánchez et al. (2021) offer a technique using a convolutional neural network (CNN) to extract features that characterize the texture, minutiae, and frequency spectrum of a fingerprint (MCC). This approach demonstrated great accuracy on the FVC2006 dataset, and the authors explicitly highlight its resilience to distortions. The essential aspect is the amalgamation of many descriptors, which allows a decrease in the incidence of false positives [5]. An analogous methodology was used in the research conducted by Anand et al. (2020), when a CNN was utilized to generate a threshold descriptor. This facilitated enhanced identification precision on the PolyU dataset [6]. Xu et al. (2010) suggested the incorporation of an ordered minutiae representation into a bit string, followed by spectral clustering. This facilitated the attainment of elevated speeds on the NIST SD14 dataset [7]. The research conducted by Yu et al. (2024) focused on the elimination of obstructive minutiae from latent fingerprints. The use of Euclidean distances enhanced identification accuracy [8]. A novel minutiae descriptor was introduced, particularly engineered for the analysis of latent fingerprints. A CNN was used for comparison, resulting in a low Equal Error Rate (EER).

Cloud platforms like as AWS are extensively used to guarantee the scalability of fingerprinting systems. Chowdhury and Imtiaz (2022) examined contact identification by deep learning techniques, attaining elevated accuracy [9]. Krishna Prakasha and Sumalatha (2025) investigated privacy in biometric systems, applicable to the incorporation of biometric identification into IoT systems using AWS Lambda, emphasizing asynchronous data processing [10]. Bortoluzzi et al. (2025) proposed a cloud-native architecture utilizing AWS, adaptable for automated fingerprint identification systems (AFIS) with DynamoDB for data caching, achieving rapid comparison speeds. They also examined the efficacy of asynchronous data processing methods in AWS for streaming, yielding a low error rate (EER) [11].

An examination of current research indicates a distinct trend towards the integration of novel fingerprint descriptors with cloud technologies. Nonetheless, there is an absence of asynchronous comparators explicitly designed for fingerprinting. Advancing innovative technologies in this domain is an urgent endeavor, facilitating both scalability and superior identification precision.

Goals and objectives. This study seeks to design and elucidate a software development technology capable of swiftly and precisely comparing fingerprints with allowable aberrations. This solution will function inside the Amazon Web Services (AWS) cloud and will not use machine learning. The primary emphasis is on velocity and the capacity to concurrently handle large quantities of data. This technique will facilitate fingerprinting and the identification of individuals using their fingerprints, particularly when substantial information requires rapid processing. To attain this objective, many activities must be undertaken:

1. Examine various fingerprint comparison techniques to identify those most appropriate for cloud implementation. Examine the use of fingerprint characteristics (minutiae), inter-point distances, and other techniques.

2. Outline the structure of the software development technology, including the necessary components for data administration, processing incoming information, and fingerprint comparison. Utilize AWS services like as S3, DynamoDB, and Lambda throughout the design process.

3. Create software that incorporates a fingerprint comparison algorithm, including possible distortions.

4. Enhance program execution efficiency by minimizing duplicate computations and using specialized libraries (NumPy, Numba).

5. Evaluate the performance using the standard FVC2000 dataset to verify its accuracy, speed, and capacity to manage substantial data quantities.

6. Establish an architecture for delivering the technology on AWS with the Cloud Development Kit (CDK) in .NET, including a Dockerfile for constructing Lambda images and interaction with Amazon Elastic Container Registry (ECR).

Asynchronous scalable comparator software development technology. A software development technology for an asynchronous, scalable fingerprint sample comparator has been created to address contemporary challenges in biometric identification, particularly in fingerprint and security systems. Current automatic fingerprint identification systems (AFIS) have difficulties in processing the increasing amounts of biometric data in real-time. This constrains their use in situations necessitating rapidity and scalability. Following a review of contemporary research in biometric identification, the following essential needs for the evolving technology were established:

- The system must function asynchronously, according to an event-driven approach to minimize data processing delays. This is essential for the effective facilitation of streaming situations, such as the real-time addition of fresh samples to the database.
- The system must possess the capability to perform thousands of comparisons per minute, autonomously adjusting to fluctuating workloads. This obviates the need for manual server resource management.
- Distortion Resistance: The system must exhibit invariance to affine transformations (translation, rotation, scaling) without relying on resource-intensive machine learning techniques. This lowers computing expenses and streamlines the implementation procedure.
- Accuracy: The system must attain elevated comparison accuracy on benchmark datasets, including FVC2000. Particular emphasis is placed on the precision of group comparisons (where a single fingerprint is evaluated against a collection of other fingerprints).
- Cloud Services Integration: The solution must exhibit complete compatibility with the AWS cloud architecture. This facilitates minimal operational expenses (pay-as-you-go) and automated resource administration.

The selection of a serverless architecture and descriptors reliant on Euclidean distance computation, excluding machine learning, was motivated by the need to save expenses, provide resilience to distortions, and facilitate asynchronous processing. This is especially crucial for fingerprint applications, since the speed and precision of recognition directly influence the dependability of the outcomes. The solution created is based on serverless

computing concepts and an event-driven approach, ensuring significant asynchronicity and scalability. The system architecture has three primary components:

• Dataset: Uploading and categorizing reference fingerprints in cloud storage, thereafter recording them in the database.

• Input Images: Processing new samples and concurrently initiating the matching procedure.

• Comparator (Matching): Evaluating fresh samples against reference fingerprint groups, consolidating outcomes, and determining matches according to established criteria.

The architectural choices were determined by the below rationale:

• Asynchronous: Employing events (S3 Events, DynamoDB Streams) to activate components eliminates continuous waiting and diminishes latency to under 100 ms per event [12]. An alternate method using periodic polling was rejected because of increased latency and excessive resource use.

• Scalability: The serverless paradigm enables AWS to automatically grow Lambda functions (up to 1000 or more instances) and the DynamoDB database (up to 40,000 operations per second), in contrast to systems reliant on fixed servers [13].

• Distortion Resistance: Geometric alignment using center of mass calculations and Euclidean descriptors provide shift invariance independently of machine learning, in contrast to convolutional neural networks (CNNs) that need pre-training.

The interaction between components occurs via events: uploading data to S3 activates a Lambda function that records the data in DynamoDB. The DynamoDB stream then activates a Lambda function to compare fingerprints. This system establishes a closed data processing loop suited for streaming. The suggested interaction system is shown in the schematic presented in Figure 1.

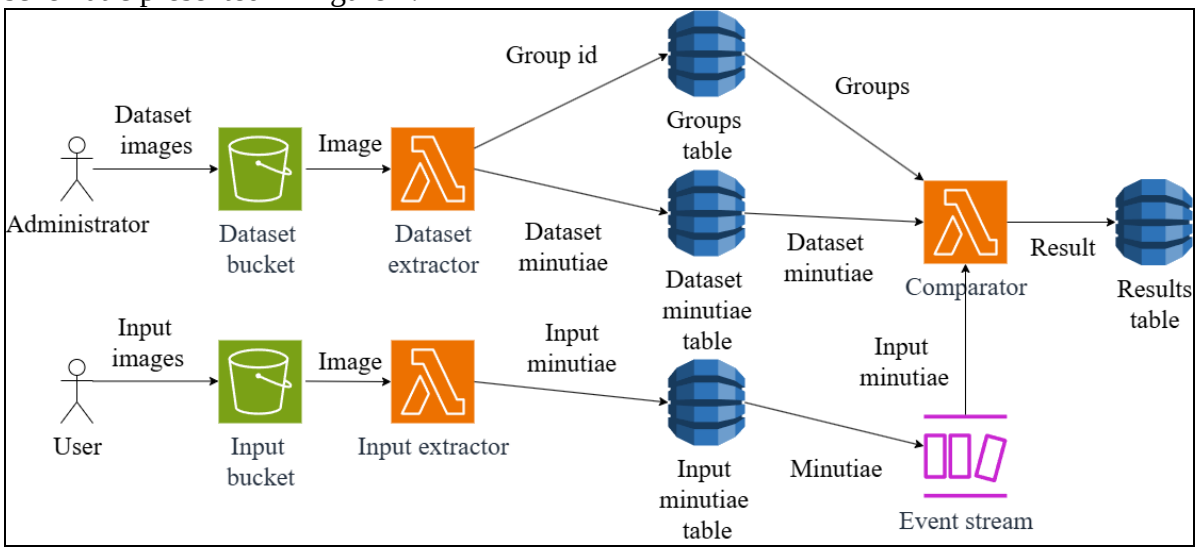


Fig. 1. Asynchronous comparator component diagram

The selection of certain AWS services was influenced by their capacity for asynchronous processing, scalability, and cost-effectiveness.

• Amazon S3 is used for the storage of original fingerprints. S3 Event Notifications are used to promptly activate a Lambda function upon the upload of fresh data, hence guaranteeing minimal latency. The benefits of S3 include boundless scalability and minimal storage expenses.

• Amazon DynamoDB: A NoSQL database used for storing metadata (ImageId as the Primary Key, GroupId as a Global Secondary Index (GSI), and metadata including center of mass coordinates (center_x, center_y) in integer format) along with comparison results (Key, Result in boolean format). Employing GSIs facilitates expedited queries on groups, whilst DynamoDB Streams provide asynchronicity.

• AWS Lambda: A serverless computing solution used for event processing, including feature extraction and fingerprint comparison. The processing algorithms are executed in Python using the NumPy and Numba libraries, while the infrastructure as code is created using .NET. Lambda may expand to 1000 or more concurrent instances and exhibits a cold start time in milliseconds when using provided concurrency.

• Amazon ECR: A repository for Lambda Docker images that includes all necessary dependencies. ECR guarantees result repeatability and compatibility with AWS CDK.

• AWS CDK: A .NET-based Infrastructure as Code (IaaC) framework for automating the deployment of cloud infrastructure, including buckets, tables, and lambdas. The CDK streamlines infrastructure administration and reduces the probability of configuration mistakes.

• AWS CloudWatch is used for monitoring slowness, failures, and cold starts. Gathering and examining logs enables the optimization of system performance and the identification of bottlenecks.

The technology design process included the following stages:

1. Data Modeling:

• Minutiae, the distinctive points of fingerprints, are represented by the MINUTIA_DTYPE_BASE structure, which includes x and y in int32 format, is_termination in boolean format, and theta in float64 format. The data format was selected for its interoperability with DynamoDB.

• Fingerprint metadata: ImageId (string), GroupId (string), center_x (integer), center_y (integer). Employing int32 for coordinates corresponds with the data format obtained from fingerprint scanners and reduces the need for conversions. Utilizing float64 would have augmented memory consumption.

2. Algorithms:

• Geometric Alignment: The minutiae coordinates are normalized with respect to the center of mass, assuring invariance to translations. The selection of the center of mass over the RANSAC method is attributed to its reduced computational cost (O(n) compared to $O(n^2)$).

• Euclidean Descriptor: Matrices of distances and angles between minutiae are computed. The resultant matrices are stored by ImageId. Threshold values for distances (7) and angles (45) were refined using the FVC2000 dataset.

• Boolean Metric Aggregation: The aggregation of comparison findings is executed with the metrics normalized_positive (average score $\geq$ 50 multiplied by the number of positive entries) and normalized_mean (overall average multiplied by the number of positive entries). The determination of fingerprint match relies on threshold values of 15.0 for normalized_positive and 1.0 for normalized_mean.

Avoiding machine learning techniques, such as convolutional neural networks (CNNs), might diminish the expenses related to training and using GPUs. Caching decreases computational duration.

3. Caching: Implementing a global cache {ImageId: {pair: value}} prevents the recalculation of metrics (distances and angles), hence decreasing the processing time for individual comparisons.

4. ID Management: The sequential assignment of IDs inside the lambda_handler (0 to N_1-1 for probes, N_1 to N_2-1 for gallery1, etc.) mitigates the risk of clashes. The method assign_unique_ids returns (array, next_start_id) to maintain consistency. The use of UUIDs was dismissed because of the enlarged key size.

5. Performance Optimization:

• Numba: Implementing JIT compilation for the greedy matching algorithm significantly enhances its speed, achieving around 0.002 seconds per call [13].

• ProcessPoolExecutor facilitates concurrent execution of local comparisons, accommodating 1000 or more cores.

• DynamoDB Global Secondary Index (GSI) facilitates rapid group query execution with millisecond response times and minimal latency in change stream processing, as seen by DynamoDB Streams latency measured in milliseconds.

Throughout the development phase, the following components were produced:

1. Dataset Element:

• Function: Uploading and categorizing reference fingerprints in DynamoDB.

• Mechanism of operation: S3 Event (object creation) triggers Lambda The function takes minutiae (e.g., from buffer), analyzes metadata, and adds data to the MinutiaeTable, using ImageId as the primary key and GroupId as the global secondary index.

Asynchronicity is accomplished by S3 Events, scalability enables the processing of thousands of files, and GSI facilitates rapid data access. The proposed alternatives (SNS/SQS) were rejected since they introduced additional delay.

2. Incoming Images Module:

• Function: Processing new samples, writing to DynamoDB, and triggering the change stream.

• Operation: S3 Event → Lambda Function → DynamoDB (Stream activates the mapping component).

3. Comparator Component:

• Function: To compare a fresh sample to reference fingerprint groups, consolidate results into metrics (normalized_pos, normalized_mea), and evaluate the match (YES/NO).

• Operational Mechanism: DynamoDB Stream → Lambda Function: ingests a new sample, searches reference fingerprint groups via GSI, assigns identifiers, verifies/calculates the cache, conducts comparisons, aggregates results, and records the conclusion.

• Illustration (Python):

```
def lambda_handler(event, context):
    dynamodb = boto3.resource('dynamodb')
    minutiae_table = os.environ['INPUT_TABLE_NAME']
    dataset_table = os.environ['DATASET_TABLE_NAME']
    group_table = os.environ['GROUP_TABLE_NAME']
    results_table = os.environ['RESULT_TABLE_NAME']
    try:
        for record in event['Records']:
            if record['eventName'] != 'INSERT':
                continue
            new_image = record['dynamodb']['NewImage']
            probe_image_id = new_image['ImageId']['S']
            probe_binary = new_image['MinutiaeBinary']['B']
            metadata = new_image['Metadata']['M']
            probe_center = (int(metadata['center_x']['N']), int(metadata['center_y']['N']))
            probe_minutiae = np.frombuffer(probe_binary,
dtype=MINUTIA_DTYPE_BASE)
            probe_minutiae, current_id = assign_unique_ids(probe_minutiae, 0)
            global_dist_cache = {}
            global_angle_cache = {}
            scan_response = minutiae_table.scan(ProjectionExpression='GroupId')
            group_ids = set(item['GroupId'] for item in scan_response['Items'] if 'GroupId'
in item)
            while 'LastEvaluatedKey' in scan_response:
                scan_response = minutiae_table.scan(ProjectionExpression='GroupId',

ExclusiveStartKey=scan_response['LastEvaluatedKey'])
```

```
                group_ids.update(item['GroupId'] for item in scan_response['Items'] if
'GroupId' in item)
            for group_id in group_ids:
                query_response =
minutiae_table.query(KeyConditionExpression=Key('GroupId').eq(group_id))
                galleries = query_response['Items']
                while 'LastEvaluatedKey' in query_response:
                    query_response =
minutiae_table.query(KeyConditionExpression=Key('GroupId').eq(group_id),

ExclusiveStartKey=query_response['LastEvaluatedKey'])
                    galleries.extend(query_response['Items'])
                group_scores = []
                group_size = 0
                for gallery in galleries:
                    if gallery['ImageId'] == probe_image_id:
                        continue
                    group_size += 1
                    gallery_image_id = gallery['ImageId']
                    gallery_binary = gallery['MinutiaeBinary']['B']
                    gallery_metadata = gallery['Metadata']['M']
                    gallery_center = (int(gallery_metadata['center_x']['N']),
int(gallery_metadata['center_y']['N']))
                    gallery_minutiae = np.frombuffer(gallery_binary,
dtype=MINUTIA_DTYPE_BASE)
                    gallery_minutiae, current_id = assign_unique_ids(gallery_minutiae,
current_id)
                    score, global_dist_cache, global_angle_cache = compare_images(
                        probe_minutiae=probe_minutiae,
                        probe_center=probe_center,
                        probe_image_id=probe_image_id,
                        gallery_minutiae=gallery_minutiae,
                        gallery_center=gallery_center,
                        gallery_image_id=gallery_image_id,
                        global_dist_cache=global_dist_cache,
                        global_angle_cache=global_angle_cache
                    )
                    group_scores.append(score)
                normalized_pos, normalized_mea = aggregate_group_scores(group_scores,
group_size)
                result = evaluate_thresholds(normalized_pos, normalized_mea)
                results_table.put_item(Item={
                    'ImageId': probe_image_id,
                    'GroupId': group_id,
                    'Result': result
                })
            minutiae_table.delete_item(Key={'ImageId': probe_image_id})
        except Exception as e:
            print(f"Error processing event: {str(e)}")
            return {'statusCode': 500, 'body': f"Error: {str(e)}"}
        return {'statusCode': 200}
```

4. Infrastructure (CDK):

- Purpose: Automation of infrastructure deployment (S3, DynamoDB, Lambda, ECR).
- Operational principle: CDK on .NET generates Docker images, uploads them to ECR, and provides the necessary resources.

The advanced technique addresses critical challenges in fingerprinting: effective handling of substantial data quantities, resilience to distortions, and elevated operational speed. The characteristics of asynchronicity and scalability provide it an appropriate option for fingerprint applications necessitating fast searches of large databases. Eschewing machine learning may diminish expenses and streamline the installation procedure. Utilizing cloud infrastructure offers reduced operational expenses and more flexibility. An exhaustive illustration of the infrastructure and technologies is presented in source [15].

Verification and outcomes of the experiment. Experiments were performed to assess the efficacy of the created asynchronous scalable fingerprint sample comparator. The standard FVC2000 dataset was used, including fingerprint photos of diverse quality and various distortions. The objective of the studies was to assess the data processing velocity, the accuracy of match identification by the system, and the scalability of the system while using AWS cloud infrastructure.

The methodology used was as follows. Testing was performed on a serverless architecture, using Amazon S3, DynamoDB, and Lambda services. The FVC2000 dataset was stored on Amazon S3. The photos underwent pre-processing to emphasize minutiae (distinctive points). This was accomplished using the methods outlined in the preceding section. Euclidean descriptors (matrices representing the distances and angles between minutiae) were computed for each picture. This information was retained in DynamoDB to expedite future comparisons. The matching technique included juxtaposing a single picture (probe) with reference prints (gallery) within the collection, in addition to doing pairwise comparisons across all photos (80 images versus 80). Cold beginnings of Lambda functions were considered during testing, since they might influence the total processing time.

The experimental findings indicated the following. The suggested asynchronous architecture exhibited its superiority throughout the data preparation phase. For instance, in provisioned capacity mode without a cold start, the processing time for the whole dataset was around 1.5 seconds, however with a cold start, the performance fell to 2 seconds for 80 photos. Secondly, the duration required to locate a single picture inside the dataset was roughly 5 seconds. This signifies that, notwithstanding the necessary number of computations, data processing transpires rapidly due to the use of caching and efficient methods using NumPy and Numba. The comprehensive comparison of all photos (80 images vs 80) required up to 15 seconds, specifically owing to horizontal scaling. The duration was contingent upon the effects of cold beginnings in the Lambda function and the process of recording results in the DynamoDB database. Cold starts increased the initial delay; however, with the implementation of provided concurrency, the processing time steadied and neared the lower threshold (about 10 seconds).

The system's accuracy was assessed using the F0.5 measure, which mostly emphasizes precision, a critical factor for fingerprinting jobs. The system attained an F0.5 score of 93% on the FVC2000 dataset. This satisfies the criteria for applied fingerprinting. The acquired data verifies that the system is resilient to affine distortions (translations, rotations, scaling) without using machine learning techniques. This facilitates a decrease in computational expenses relative to methods using CNNs.

Scenarios with escalating load (up to 1000 photos in 10 groups) were developed to evaluate the system's scalability. The serverless design and dynamic scalability of Lambda and DynamoDB enabled the system to effectively execute thousands of comparisons per minute. Manual resource management was unnecessary. DynamoDB Streams with S3 Events facilitated low-latency (in milliseconds) asynchronous processing.

Conclusions. The advanced software development technology for fingerprint comparison on the AWS cloud allows rapid and precise biometric identification without the need of machine

learning. Experiments using the FVC2000 database indicated that retrieving a single picture takes around 5 seconds, while a comprehensive comparison of 80x80 photos necessitates up to 15 seconds owing to horizontal scaling, achieving an F0.5 score of 93%.

The system utilizes a serverless architecture including Amazon S3, DynamoDB, and Lambda, enabling it to do thousands of comparisons per minute with little latency and automated scalability. Utilizing caching and JIT compilation (Numba) accelerates computations, while minimizing machine learning applications decreases expenses.

This method is appropriate for fingerprint recognition systems where rapidity and precision are essential while handling large datasets. The utilization of AWS cloud diminishes expenses and enhances system adaptability, rendering it a compelling subject for future investigation and use in fingerprinting.

References

1. Yin X., Zhu Y., Hu J. A Survey on 2D and 3D Contactless Fingerprint Biometrics: A Taxonomy, Review, and Future Directions. *IEEE Open Journal of the Computer Society*. 2021. Vol. 2. P. 370–381. DOI: 10.1109/OJCS.2021.3119572.
2. Siddiqui M., Iqbal S., Al-Haqbani B., Al-Shammari B., Khan №., Razzak I. A Robust Algorithm for Contactless Fingerprint Enhancement and Matching. 2024 *International Conference on Digital Image Computing: Techniques and Applications (DICTA). Perth, Australia*. 2024. P. 214–220. DOI: 10.1109/DICTA63115.2024.00041.
3. Herbadji A., Guermat N., Akhtar Z. Deep neural networks based contactless fingerprint recognition. *2nd International Conference on New Technologies of Information and Communication (NTIC). Mila, Algeria*. 2022. P. 1–6. DOI: 10.1109/NTIC55069.2022.10100455.
4. Irshad R.R., Hussain S., Hussain I., Nasir J.A., Zeb A., Alalayah K.M. IoT-Enabled Secure and Scalable Cloud Architecture for Multi-User Systems: A Hybrid Post-Quantum Cryptographic and Blockchain-Based Approach Toward a Trustworthy Cloud Computing. *IEEE Access*. 2023. Vol. 11. Pp. 105479–105498. DOI: 10.1109/ACCESS.2023.3318755.
5. Pérez-Sánchez I., Cervantes B., Medina-Pérez M.A., Monroy R., Loyola-Gonzáles O., García S. An Indexing Algorithm Based on Clustering of Minutia Cylinder Codes for Fast Latent Fingerprint Identification. *IEEE Access.* 2021. Vol. 9. Pp. 85488–85499. DOI: 10.1109/ACCESS.2021.3088314.
6. Anand V., Kanhangad V. PoreNet: CNN-Based Pore Descriptor for High-Resolution Fingerprint Recognition. *IEEE Sensors Journal*. 2020. Vol. 20, No. 16. P. 9305–9313. DOI: 10.1109/JSEN.2020.2987287.
7. Xu H., Veldhuis R.N.J. Spectral Minutiae Representations for Fingerprint Recognition. *Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing. Darmstadt, Germany*. 2010. P. 341–345. DOI: 10.1109/IIHMSP.2010.90.
8. Yu J., Niu L., Gao C., Cao Z., Zhao H. Partial Fingerprint Matching via Feature Similarity and Pre-training. *IEEE International Joint Conference on Biometrics (IJCB). Buffalo, NY, USA.* 2024. P. 1–9. DOI: 10.1109/IJCB62174.2024.10744474.
9. Chowdhury A.M.M., Imtiaz M.H. Contactless Fingerprint Recognition Using Deep Learning — A Systematic Review. *Journal of Cybersecurity and Privacy*. 2022. Vol. 2, No. 3. P. 714–730. DOI: 10.3390/jcp2030036.
10. Krishna Prakasha K., Sumalatha U. Privacy-Preserving Techniques in Biometric Systems: Approaches and Challenges. *IEEE Access*. 2025. Vol. 13. P. 32584–32616. DOI: 10.1109/ACCESS.2025.3541649.
11. Bortoluzzi F., Irwin B., Westphall C.M. Cloud Telescope: An Ephemeral, Distributed, and Cloud-Native Architecture for Collecting Internet Background Radiation. *IEEE Access*. 2025. Vol. 13. P. 45682–45714. DOI: 10.1109/ACCESS.2025.3549623.

12. Alotaibi A., Hussain M., Aboalsamh H.A. Cross-Sensor Fingerprint Recognition Using Convolutional Neural Network and Canonical Correlation Analysis. *IEEE Access*. 2024. Vol. 12. P. 84738–84751. DOI: 10.1109/ACCESS.2024.3413975.
13. Sanchez-Fernandez A.J. et al. Asynchronous Processing for Latent Fingerprint Identification on Heterogeneous CPU-GPU Systems. *IEEE Access*. 2020. Vol. 8. P. 124236–124253. DOI: 10.1109/ACCESS.2020.3005476.
14. Amazon Web Services. DynamoDB Metrics and Dimensions. URL: https://docs.aws.amazon.com/amazondynamodb/latest/developerguide/metrics-dimensions.html
15. Pohuliaiev Y. Asynchronous comparator software repository. URL: https://github.com/matan4life/PHD.

ТЕХНОЛОГІЯ РОЗРОБКИ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ АСИНХРОННОГО МАСШТАБОВАНОГО КОМПАРАТОРА ДАКТИЛОСКОПІЧНИХ ЗРАЗКІВ НА БАЗІ ХМАРНОЇ ІНФРАСТРУКТУРИ

Ю. С. Погуляєв, К. С. Смеляков

Харківський національний університет радіоелектроніки
14, Науки пр., Харків, 61166, Україна
Emails: yurii.pohuliaiev@nure.ua, kyrylo.smelyakov@nure.ua

Дактилоскопія є ключовим методом біометричної ідентифікації в системах безпеки, судовій експертизі та контролі доступу завдяки унікальності папілярних узорів. Проте обробка великих обсягів даних та афінні перетворення (зсуви, повороти, масштабування) створюють проблеми для автоматизованих систем ідентифікації відбитків пальців (AFIS). Сучасні хмарні платформи, такі як AWS, пропонують рішення для масштабованості та асинхронної обробки, але їх інтеграція з дактилоскопічними компараторами потребує подальшого вивчення. Мета дослідження: запропонувати технологію розробки програмного забезпечення для асинхронного масштабованого компаратора дактилоскопічних зразків на базі AWS, що забезпечує швидку та точну ідентифікацію без використання машинного навчання та є стійкою до афінних спотворень. Була створена безсерверна архітектура з використанням AWS, включаючи S3 для зберігання, DynamoDB для метаданих та Lambda для обчислень. Для стійкості до спотворень застосовувалися евклідові дескриптори, кешування для підвищення ефективності обчислень, а також бібліотеки NumPy і Numba для оптимізації. Оцінка проводилася на наборі даних FVC2000 з використанням подієвої моделі (S3 Events, DynamoDB Streams) та .NET CDK для автоматизації інфраструктури. Експерименти на FVC2000 показали точність F0.5 = 93%, час пошуку одного зображення — близько 5 секунд, а повне порівняння 80x80 зображень — до 15 секунд. Система підтримує тисячі порівнянь за хвилину завдяки автоматичному масштабуванню Lambda та DynamoDB. Асинхронність та відмова від машинного навчання забезпечують зниження витрат і підвищення швидкості. Дослідження сприяє розвитку дактилоскопії, пропонуючи ефективний підхід до обробки великих наборів даних без ресурсоємних методів, інтегрований із хмарними сервісами AWS, що розширює можливості AFIS. Технологія може бути застосована в системах безпеки, судовій експертизі та контролі доступу, забезпечуючи швидку й точну ідентифікацію з низькими витратами завдяки хмарній інфраструктурі.

Ключові слова: дактилоскопія, хмарні обчислення, асинхронна обробка, масштабованість, AWS Lambda, евклідові дескриптори, компаратор

DOI 10.15276/imms.v15.no4.496
UDC 004.932.1

Informatics and Mathematical Methods in Simulation
Vol.15 (2025), No. 4, pp. 496-506

MODIFICATION OF THE MASK R-CNN ARCHITECTURE FOR IMAGE DETECTION AND SEGMENTATION

N. Volkova, M. Shvandt

National Odesa Polytechnic University
1, Shevchenko Ave., Odesa, Ukraine, 65044
Emails: volkova.n.p@op.edu.ua, maxim.shvandt@gmail.com

The paper addresses the problem of detecting and segmenting animal images. An analysis of neural network architectures for detecting and segmenting objects is carried out. A neural network architecture is proposed for detecting and segmenting images with a fuzzy background and partial overlapping of objects based on a modification of the Mask R-CNN architecture, which demonstrates a sufficiently high evaluation indicators of accuracy and quality of segmentation and is able to utilize additional image features of multichannel images. The main elements of the proposed architecture are a double branch of a feature extractor with feature fusion that uses additionally obtained image features. The proposed architecture was tested on a set of test images of experimental animals. The results of detecting and segmenting experimental animals by the proposed architecture and several basic Mask R-CNN variants were compared. The segmentation quality was assessed using the accuracy (Accuracy, Precision), completeness (Recall) metrics. Based on experimental studies, it was determined that training a modification of the Mask RCNN architecture for 50 epochs allows obtaining sufficiently high indicators of quality and accuracy of detection and segmentation, namely: accuracy (Accuracy) - 0.9, precision (Precision) - 0.92, completeness (Recall) - 0.92, while maintaining basic operability. Mask R-CNN variants with ResNet18/34 feature extractors have lower accuracy, and basic Mask R-CNN with ResNet50/101 have significantly larger sizes without the possibility of using additional image features. Thus, the architecture proposed in the work is effective for tasks of detection and segmentation of objects that require high accuracy and quality of their localization in the image.

Keywords: neural network; Mask R-CNN, architecture; object detection; segmentation, object tracking; evaluation indicators

Introduction. In recent decades, the systematic study of animal behavior has assumed increasing importance across disciplines including neuroscience, ethology, environmental sciences, and agricultural research. Behavioral responses to environmental stimuli provide not only key indicators of animal health and welfare but also serve as sensitive proxies for broader ecological dynamics [1-3]. Parallel to these scientific needs, advances in computational technologies, most notably in computer vision, have enabled the quantification and analysis of behavior with levels of precision, reproducibility, and scalability that were previously unattainable [4-5].

Historically, behavioral studies relied heavily on manual observation or basic sensor systems. Such approaches, while being foundational, were very labor-intensive, constrained in temporal and spatial resolution, and susceptible to observer bias [4]. The emergence of high-resolution imaging, sophisticated motion analysis, and automated tracking algorithms has transformed this landscape, allowing continuous and non-invasive monitoring of animal activity across diverse experimental and natural contexts. Although deep learning has become one of the leading tools in the field, many influential studies continue to use classic computer vision techniques such as background subtraction, contour detection, and trajectory clustering, which remain effective and interpretable for a wide range of behavioral analyses [7,8].

However, the adoption of neural networks has rapidly expanded the scope of computational ethology. Deep learning frameworks, especially, convolutional and recurrent

architectures, are increasingly used for pose estimation, fine-grained action recognition, and multi-animal tracking. These methods allow researchers to capture subtle behavioral nuances, integrate multimodal data streams, and generalize across species and contexts [9]. Toolkits such as DeepBehavior demonstrate how deep learning can be applied to both animal and human behavior imaging data, providing accessible pipelines for neuroscience and ethology [9]. Recent comprehensive studies have highlighted the breadth of applications of deep learning in the study of animal behavior, ranging from bioacoustics to video tracking, and have outlined challenges and opportunities for future research [10]. Neural networks are also being leveraged to link behavioral signatures with physiological or genetic data, offering a powerful bridge between observable actions and underlying biological mechanisms.

The growing popularity of automated behavioral analysis is driven by a combination of technological readiness and scientific necessity. Global challenges such as climate change, disease emergence, and food security demand real-time tools to assess animal responses to environmental pressures [11]. Moreover, behavior often reflects underlying physiological or cognitive states more rapidly than biochemical markers can detect. Automated behavioral metrics therefore provide a non-invasive view into animal welfare and can even serve as early -warning systems for stress or illness [12,13]. Together, these developments underscore the central role of computational methods, ranging from traditional vision algorithms to state-of-the-art neural networks, in shaping the future of animal behavior research.

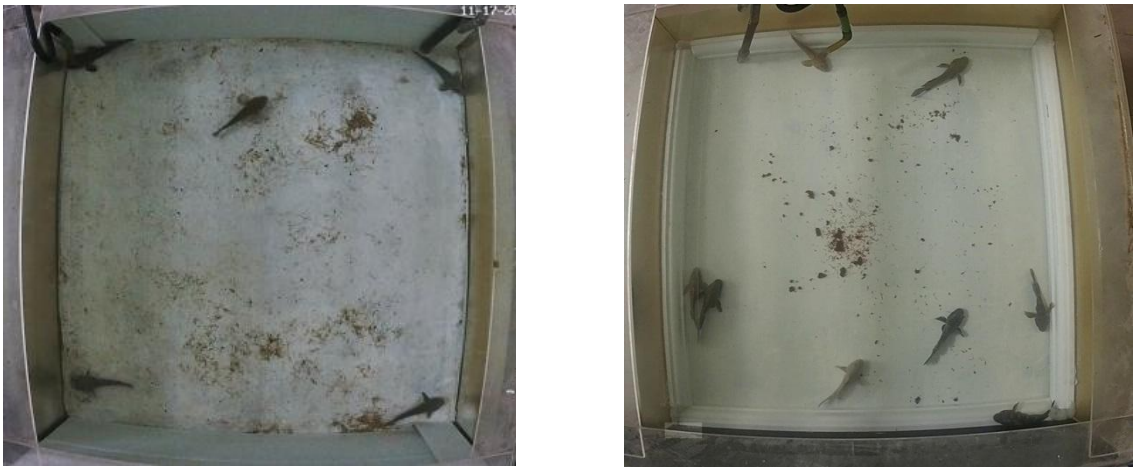
Analysis of recent publications and formulation of the problem. As noted earlier, mice and fish are among the most commonly employed model organisms in ecotoxicology and ethology. A substantial portion of such research relies on the study of animal behavior under controlled, enclosed conditions, particularly during the initial stages of experimentation. For rodents, the test setup involves a perforated test box designed to stimulate exploratory activity. A fixed overhead camera is positioned above the apparatus to continuously record the animals over extended periods, ranging from 30 minutes to several hours. Behavioral observations in this context often focus on quantifying movements between holes and documenting investigatory attempts to peer into them, which are interpreted as indicators of exploratory drive and territorial assessment (Fig. 1a–1b).

a *b*

Fig. 1. Mice/rats behaviour study: *a* – screenshot of a video with a lab rat, *b* – screenshot of a video with lab mice

In our particular case the behavior of fish (especially bullheads) is of considerable interest to the researchers. The enclosed test environment is represented by a square aquarium filled with real sea water to simulate the real environmental conditions. The aquarium is also supplied with oxygen via pipes, and a camera is mounted above the aquarium and records the bullheads over a period of time (from several hours to one day). The video recording is divided into 30-minute video segments to facilitate further analysis (Fig. 2a-2b).

In terms of the experiment here are up to 10 subjects supposed to be placed inside the aquarium. The behaviour patterns that currently are of greatest interest to the researchers among other ones are: the general number of movements (position changes) for each subject; the general number of movements for complete test group; number of attacks/conflicts between 2 subjects (a movement of subject A towards subject B that results in subject B escaping in other direction); keeping each object`s ID during the complete recording session. This is particularly important in case of placement of 2 or 3 different kinds of bullheads into the aquarium.



a *b*

Fig. 2. Fish (bullheads) behaviour study: *a, b* – screenshots from videos with bullheads in the test environment (aquarium)

Currently, analysis of video sequences is performed manually by laboratory personnel, which is a time-consuming and insufficiently accurate process, since behavioral patterns are formed subjectively. This highlights the need for an automated algorithm capable of detecting, tracking, and analysing subject movements with greater efficiency and precision. Developing such a algorithm, however, is complicated by several factors. Although the position of a subject typically does not shift dramatically between two or three consecutive frames, its direction of movement can change abruptly, making trajectory estimation challenging. In addition, while the experimental environment is generally stable, the aquarium background is subject to dynamic changes. Food can introduced before and during test sessions, and the by-products of bullhead activity accumulate on the aquarium floor. These sediments often migrate in the form of clumps under the influence of water currents generated by swimming fish and aeration from oxygen tubes. Because these clumps may resemble the fish in colour and can grow up to considerable size, distinguishing the animals when they swim above such deposits becomes difficult. This variability also precludes the use of simple strategies such as colour-based tracking. Therefore, all these aspects must be taken into account when developing a robust and reliable tracking algorithm.

Another important problem during detection and tracking before performing object identification is the separation of objects that may merge into one spot when approaching each other. That is why it was decided to add a neural detector capable of segmenting objects to the proposed object detection and tracking algorithm, along with basic detection methods such as background subtraction [14].

Object detection with segmentation, often referred to as instance segmentation, is a central task in computer vision. Unlike pure object detection, which predicts bounding boxes and class labels, instance segmentation requires delineating each object at the pixel level. This dual requirement makes the task more challenging but also more useful in domains such as

autonomous driving, medical imaging, and video surveillance. Over the past decade, researchers have proposed a variety of neural network architectures to address this problem. This overview examines four representative families: Mask R‑CNN, YOLACT, SOLO/SOLOv2, and DetectoRS, while also briefly noting related approaches such as CondInst and Panoptic FPN. Each architecture embodies different trade‑offs between accuracy, speed, and complexity.

Mask R-CNN [15] is the most widely recognized architecture for instance segmentation. It extends Faster R-CNN [16], a two-stage object detector, by adding a parallel branch dedicated to mask prediction. The model begins with a convolutional backbone, typically ResNet-50 or ResNet-101, augmented with a Feature Pyramid Network (FPN) [17] to capture multi-scale features. A Region Proposal Network (RPN) then generates candidate object regions, which are refined through RoIAlign, a pooling method that avoids the misalignments caused by quantization in earlier designs. For each region of interest, the network predicts class labels, bounding box coordinates, and a binary mask. This modular design allows Mask R-CNN to achieve high accuracy across benchmarks such as COCO, and its flexibility has made it a standard baseline in both research and industry. However, the two-stage nature of the model makes it computationally heavy and relatively slow, which can be a limitation in real-time applications. In contrast to the two-stage paradigm, YOLACT [18] demonstrates that instance segmentation can be achieved in real time. Instead of predicting masks for each region of interest, YOLACT generates a set of prototype masks for the entire image and then linearly combines them with per-instance coefficients predicted by the detection branch. This one-stage design is conceptually simpler and avoids the overhead of region proposals and RoI operations. The result is a system that can run at real-time speeds on modern GPUs, making it attractive for applications such as robotics or video analysis where latency is critical. The trade-off is that YOLACT generally achieves lower accuracy than two-stage methods, particularly on small or overlapping objects, and its prototype mask approximation can blur fine boundaries. Nevertheless, it still represents a step toward balancing speed and accuracy in instance segmentation.

SOLO (Segmenting Objects by Locations) [19] and its successor SOLOv2 [20] take a different approach by reformulating instance segmentation as a direct classification problem on a spatial grid. The image is divided into grids, and each grid cell predicts whether it belongs to an object instance and outputs a mask kernel to generate the segmentation. This anchor‑free, grid‑based formulation eliminates the need for proposals or RoI pooling, making the pipeline more straightforward. SOLOv2 improves upon the original by introducing dynamic convolution, which enhances both speed and accuracy. These models demonstrate that instance segmentation can be addressed in an end‑to‑end manner without the complexities of proposal generation. However, the grid‑based formulation can be sensitive to object scale and placement, and dense scenes with many overlapping objects remain challenging. Despite these limitations, SOLO and SOLOv2 highlight the potential of anchor‑free methods and have influenced subsequent research. As an alternative, DetectoRS [21] represents the high‑accuracy, high‑complexity end of the spectrum. Built upon Cascade Mask R‑CNN [22], it introduces two key innovations: Recursive Feature Pyramid (RFP) and Switchable Atrous Convolution (SAC). RFP enhances multi‑scale feature representation by feeding FPN outputs back into the backbone, allowing recursive refinement of features. SAC adapts receptive fields dynamically by switching between different atrous rates, enabling the network to capture varied contextual information. Together, these innovations significantly improve performance, and DetectoRS has achieved state‑of‑the‑art accuracy on COCO

benchmarks. The cost of this performance is computational expense: the model is very resource‑intensive, requires significant GPU memory, and is unsuitable for real‑time applications. Its complexity also makes it harder to train and deploy compared to simpler designs. Nonetheless, DetectoRS demonstrates how architectural innovations can push the boundaries of accuracy in instance segmentation.

Instance segmentation has progressed rapidly, moving from the proposal‑based, two‑stage paradigm of Mask R‑CNN to real‑time one‑stage models like YOLACT, to anchor‑free formulations like SOLOv2, and to advanced recursive designs like DetectoRS. Each architecture reflects a different balance between accuracy, speed, and complexity, and each has influenced subsequent research. The diversity of approaches underscores the richness of the problem and the ongoing search for architectures that can deliver both precision and efficiency. As applications of computer vision expand, the demand for models that can perform accurate, real‑time instance segmentation will continue to drive innovation, likely leading to architectures that combine the strengths of current paradigms while mitigating their weaknesses.

The considered models are initially designed for typical images with 3 channels without additional features. Also, not all architectures have a flexible modular structure. Therefore, the goal of the study is to create an architecture that could use the advantages of the presence of additional image features while maintaining the overall accuracy of the basic options with a smaller structure/size of the model (compared, for example, to the basic Mask R-CNN with ResNet101) along with preserving the relative modularity for further modifications if necessary. To achieve the set goal, it is necessary to solve the following main tasks: analysis of the architectures of object detection and segmentation networks; development of the basis of a feature extractor that will use additional image feature channels; experimental study of the developed architecture.

Main material. Image processing within the proposed approach [14] includes several sequential stages: preprocessing, primary segmentation (for further background filtering), detection, and analysis. During image preprocessing and earlier detection stages, such as background subtraction [14], we are able to obtain additional image features in form of new image channels, such as the result of background subtraction/image with CLAHE or shadow correction. These channels contain additional image representations with highlighting specific image aspects. Thus it was decided to utilize these additional channels to enhance object detection and segmentation in terms of combined detector. The basic Mask R-CNN was chosen for it general accuracy and primarily, for its modularity since currently it is not required to adopt the detection for real-time usage. Also there was an aim to create an architecture that with the utilization of additional channels would be smaller in terms of basic Mask R-CNN with ResNet101 backbone or even ResNet50.

Since using a separate backbone for additional image data is already a valid approach for extra data handling [23,24], it was decided to use to parallel feature extraction branches in the backbone, one for main image channels (Branch A) and one for feature (engineered) channels (Branch B). Since additional channels are refined and filtered from noise, to enhance feature constructions the feature fusion was applied from branch B into branch A. Among different fusion technics [25,26] it was currently decided to utilize SE-based fusion [27] since using it one will stick to data-aware mixing: after concatenation, the SE gate learns which channels (from A or B) are useful per stage and per image, instead of blindly summing. Also it is robust to scale/statistics mismatch as the channel attention can down-weight noisy or poorly scaled auxiliary features (e.g., depth/engineered bands) before they enter the top-down pathway and is Lightweight & stable, as the squeeze-excite MLP adds very few parameters/compute and preserves spatial structure, which is good for plugging into FPN

stages without upsetting optimization. Table 1 and Figure 3 show the general idea of merging two branches within the backbone.

The SE-based fusion performs:

a) concatenation of features from the main (A) and aux (B) branches at the same spatial scale:

$$U = \mathrm{Concat}\left(A, B\right) \in \mathrm{R}^{H\times W\times C},\ C = C_A + C_B,$$

where *H,W,C* are height, width and number of channels.

b) feature squeezing (global context per channel):

$$z_c = \frac{1}{HW}\sum_{i=1}^{H}\sum_{j=1}^{W} U_{i,j,c} \Rightarrow z \in \mathrm{R}^{C};$$

c) learning of channel importances with a bottleneck set by *r* (SE-ratio)

$$s = \sigma\left(W_2\delta\left(W_1 z\right)\right),\ W_1 \in \mathrm{R}^{\frac{C}{r}\times C},\ W_2 \in \mathrm{R}^{C\times\frac{C}{r}},$$

where *r* is SE-ratio, $\delta = \mathrm{ReLU}$, $\sigma = \mathrm{Sigmoid}$.

d) channels reweighting:

$$\tilde{U}_{i,j,c} = s_c \cdot U_{i,j,c}.$$

Table 1.

General dual-branch fused feature extractor backbone

Stage	Spatial stride	A-branch (main)	B-branch (aux)	Fusion points (B→A)	Output for FPN
Input	1×	Split 6-ch → A: ch0–2, B: ch3–5	same	–	–
C1 (stem)	2×	conv7×7/2, BN, ReLU → 64ch; maxpool/2	same	–	–
C2 (layer1)	4×	ResNet-18 layer1 → 64ch	64ch	–	C2(A) → goes to FPN lateral 1×1
C3 (layer2)	8×	128ch	128ch	Stage-entry fuse at C3: B→A (se_concat)	C3(A′) → FPN lateral
C4 (layer3)	16×	256ch	256ch	Blockwise fuse inside C4 (after blocks)	C4(A′) → FPN lateral
C5 (layer4)	32×	512ch	512ch	Blockwise fuse inside C5 (after blocks)	C5(A′) → FPN lateral

Research results. To assess overall performance and applicability in more realistic conditions, all training and testing sessions were conducted on locally available hardware. Test experiments were conducted on a laptop with an Intel Core i9-13980HX CPU, 64GB RAM and a single NVidia GeForce RTX 4090 Laptop GPU. As a relative primary test the basic Mask RCNN variant with several single backbones and training parameters were trained and tested. They include the standard configuration with ResNet50 and ResNet101 backbones, as well as a version with smaller backbone, such as Resnet18 and ResNet34. Due to GPU memory limitations the batch size has to be altered in order to perform training on available hardware. The number of epochs was fixed at value when model accuracy stopped increasing significantly.

For basic Mask RCNN architectures a dataset with JPG-images was used. It was separated into 340 train and 13 validation images (the dataset was augmented using image rotations). The example benchmarks are presented on table 2. The architecture implementation from [28,29] was used as an additional source of implementation example.

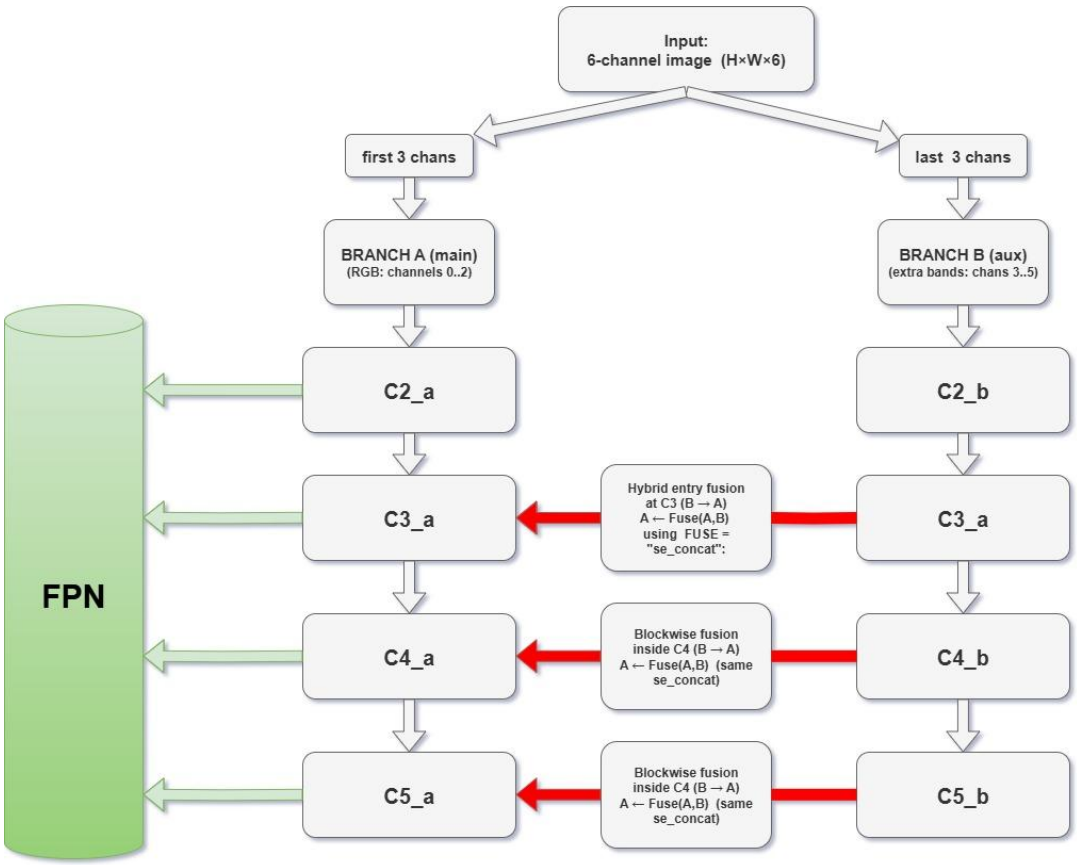


Fig. 3. Visualization of the general scheme of fusion of two branches inside the backbone

The elements of the confusion matrix were used to calculate segmentation quality scores: Accuracy (1), Precision (2), Recall (3) [30]:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \tag{1}$$

$$\text{Precision} = \frac{TP}{TP + FP}, \tag{2}$$

$$\text{Recall} = \frac{TP}{TP + FN}. \tag{3}$$

Table 2.

Basic Mask R-CNN architecture training and benchmarking

Architecture	Backbone	Batch size	Training epochs	mAP50	AVG Inference ms (excluding model 1st sample warm-up)	Precision	Recall
Mask R-CNN	ResNet101	2	30	0.92	102.5	0.963	0.939
Mask R-CNN	ResNet50	4	30	0.92	101.8	0.974	0.927
Mask R-CNN	ResNet34	6	30	0.84	~91	0.89	0.89
Mask R-CNN	ResNet18	4	30	0.83	104.9	0.86	0.902

For the developed dual-branch feature extractor architecture, the number of training epochs was increased to 50. The benchmark is shown on table 3. Figure 3 shows the example detection result.

Table 3.

Dual Fused ResNet18 backbone Mask R-CNN architecture benchmarking

Architecture	Backbone	Batch size	Training epochs	mAP50	AVG Inference ms (excluding model 1st sample warm-up)	Precision	Recall
Dual Branch Fused backbone Mask R-CNN	ResNet18	4	50	0.9	105.5	0.915	0.915

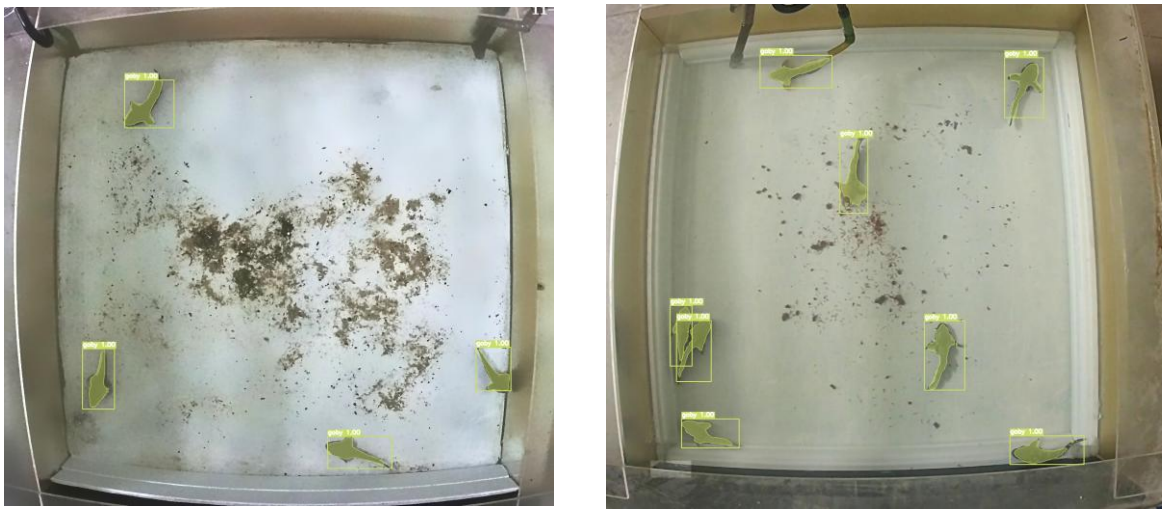


Fig. 3. Dual fused backbone detection visualization

Conclusions. This research proposes a neural network architecture for detecting and segmenting images with a fuzzy background and partial overlapping of objects based on a modification of the Mask R-CNN architecture. The accuracy and quality of object detection and segmentation in images using the proposed modified architecture of the Mask R-CNN neural network with a dual feature extractor on multi-channel images were studied. To increase the volume of the training sample, the paper performed a data augmentation procedure, the volume of which was increased by 4 times. The segmentation quality was assessed using an confusion matrix, based on the elements of which segmentation accuracy indicators (Accuracy, Precision, Recall) were calculated. Comparative analysis showed that the proposed modified architecture of the Mask R-CNN network with a smaller number of parameters compared to the variants with one ResNet101/ResNet50 shows accuracy and speed at the level of these architectures. Also, the indicators of the modified architecture exceed the accuracy indicators for the variants with ResNet18/34. The presence of an additional branch for additional features also allows working with images with more than 3 channels for a wider use of all image features. Further research will be aimed at additional improving of the architecture accuracy and to execution of object identification.

Akwnoledgments. Special thanks for the research assistance and provided test videos and images of lab animals to Faculty of Biology of Odesa I.I. Mechnikov National University.

References

1. Tinbergen N. On aims and methods of ethology. *Zeitschrift für Tierpsychologie*. 1963. 20(4). 410‑433. URL:https://doi.org/10.1111/j.1439-0310.1963.tb01161.x

2. Dawkins M. Behavior as a tool in welfare assessment. *Applied Animal Behaviour Science*. 2004. V.86(3-4). P.227-233. URL:https://doi.org/10.1016/j.applanim.2004.02.001
3. Guo C., Chen Y., Ma C., Hao S., Song J. A survey on AI-driven mouse behavior analysis applications and solutions. *Bioengineering*. 2024. V.11(11). P.1121. URL:https://doi.org/10.3390/bioengineering11111121
4. Feng J.-X., Li P., Liu Y., Liu L., Li Z.H. A latest progress in the study of fish behavior: Cross-generational effects of behavior under pollution pressure and new technologies for behavior monitoring. *Environmental Science and Pollution Research*. 2024. V.31. P. 11529-11542. URL:https://doi.org/10.1007/s11356-024-31885-2
5. Dell A., Bender J., Branson K., Couzin I., Polavieja G. Automated image-based tracking and its application in ecology. *Trends in Ecology & Evolution*. 2014. V.29(7). P.417-428. URL:https://doi.org/10.1016/j.tree.2014.05.004
6. Anderson, D., Perona, P. Toward a science of computational ethology. *Neuron*. 2014. V.84(1). 18-31. https://doi.org/10.1016/j.neuron.2014.09.005
7. Yin Z., Xiao L., Ma R., Han Z., Li Y. Detecting abnormal animal behaviors using optical flow and background subtraction. *Computers and Electronics in Agriculture*. 2020. 174. P.105471. URL:https://doi.org/10.1016/j.compag.2020.105471
8. Beyan C., Fisher R. Animal behavior recognition using spatio-temporal features. *Pattern Recognition*. 2018. No.76. P.12-22. URL:https://doi.org/10.1016/j.patcog.2017.10.008
9. Arac A., Zhao P., Dobkin B. H., Carmichael S. T., Golshani P. DeepBehavior: A deep learning toolbox for automated analysis of animal and human behavior imaging data. *Frontiers in Systems Neuroscience*. 2019. No. 13. P. 20. URL: https://doi.org/10.3389/fnsys.2019.00020
10. Fazzari E., Romano D., Falchi F., Stefanini C. Animal behavior analysis methods using deep learning: A survey. *arXiv preprint*. 2023. URL: https://arxiv.org/abs/2405.14002
11. Manteuffel G., Puppe B., Schön P., Bruckmaier R., Janssen D. Sensor-based analysis of animal behavior. *Animal*. 2009. No.3(9). P.1197 - 1204. URL: https://doi.org/10.1017/S1751731109004526
12. Neethirajan S. Recent advances in wearable sensors for animal health management. *Sensors and Biosensors Research*. 2017. No.20. P.1 - 11. URL: https://doi.org/10.1016/j.sbsr.2018.02.004
13. Spampinato, C., Palazzo, S., Boom, B., Lin, H., Wei, J., et al. Understanding fish behavior during typhoon events in real-life underwater environments. *Multimedia Tools and Applications*. 2014. 70(1). 199-236. https://doi.org/10.1007/s11042-012-1101-5
14. Volkova, M., Shvandt. Segmentation-based approach for object detection. *Proceedings of Odessa Polytechnic University*. 2025. 1(71). P. 145 - 156. https://doi.org/10.15276/opu.1.71.2025
15. He K., Gkioxari G., Dollár P., Girshick R. Mask R-CNN. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. 2017. URL: https://doi.org/10.48550/arXiv.1703.06870
16. Ren S., He K., Girshick R., Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems (NeurIPS)*. 2015. URL: https://doi.org/10.48550/arXiv.1506.01497

17. Lin T.Y., Dollár P., Girshick R., He K., Hariharan B., Belongie S. Feature pyramid networks for object detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017. URL: https://doi.org/10.48550/arXiv.1612.03144
18. Bolya D., Zhou C., Xiao F., Lee Y.J. YOLACT: Real‑time instance segmentation. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. 2019. URL: https://doi.org/10.48550/arXiv.1904.02689
19. Wang X., Kong T., Shen C., Jiang Y., Li L. SOLO: Segmenting objects by locations. *Proceedings of the European Conference on Computer Vision (ECCV)*. 2020. URL: https://doi.org/10.48550/arXiv.1912.04488
20. Wang X., Zhang R., Kong T., Li L., Shen C. SOLOv2: Dynamic and fast instance segmentation. *Advances in Neural Information Processing Systems (NeurIPS)*. 2020. URL: https://doi.org/10.48550/arXiv.2003.10152
21. Qiao S., Chen L.C., Yuille A. DetectoRS: Detecting objects with recursive feature pyramid and switchable atrous convolution. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2021. URL: https://doi.org/10.48550/arXiv.2006.02334
22. Cai Z., Vasconcelos N. Cascade R‑CNN: Delving into high quality object detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2018. URL: https://doi.org/10.48550/arXiv.1712.00726
23. Gupta S., Girshick R., Arbelaez P., Malik J. Learning rich features from RGB‑D images for object detection and segmentation. *European Conference on Computer Vision (ECCV)*. 2014. P. 345‑360. URL: https://doi.org/10.1007/978-3-319-10584-0_23
24. Zhang Z., Zhang J., Bailo O., Vázquez D., Xu J., López A.M. A two‑branch feature fusion framework for RGB‑D pedestrian detection. *Remote Sensing*. 2022. No.14(3). P.645. URL: https://doi.org/10.3390/rs14030645
25. Woo S., Park J., Lee J.Y., Kweon I. S. CBAM: Convolutional block attention module. *European Conference on Computer Vision (ECCV)*. 2018. P.3 ‑ 19. URL: https://doi.org/10.1007/978-3-030-01234-2_1
26. Perez E., Strub F., de Vries H., Dumoulin V., Courville A. FiLM: Visual reasoning with a general conditioning layer. *International Conference on Learning Representations (ICLR)*. 2018. URL: https://doi.org/10.48550/arXiv.1709.07871
27. Hu J., Shen L., Sun G. Squeeze‑and‑excitation networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2018. P. 7132‑7141. URL: https://doi.org/10.1109/CVPR.2018.00745
28. Mask R‑CNN for object detection and instance segmentation on Keras and TensorFlow. *GitHub repository*. 2017. URL: https://github.com/matterport/Mask_RCNN
29. Waleed A., Mask R. CNN: Train on the Balloon Dataset and Run Color Splash. *Medium*. 2018. URL: https://engineering.matterport.com/splash-of-color-instance-segmentation-with-mask-r-cnn-and-tensorflow-7c761e238b46
30. Sathyanarayanan S., Roopashri-Tantri B. Confusion Matrix ‑ Based Performance Evaluation Metrics. *African Journal of Biomedical Research*, 2024. No.27(4S). P.4023-4031. DOI: 10.53555/AJBR.v27i4S.4345

N. Volkova, M. Shvandt

МОДИФІКАЦІЯ АРХІТЕКТУРИ MASK R-CNN ДЛЯ ДЕТЕКТУВАННЯ ТА СЕГМЕНТАЦІЇ ЗОБРАЖЕНЬ

Н. Волкова, М. Швандт

Національний університет «Одеська політехніка»
1, Шевченка пр., Одеса, 65044, Україна
Emails: volkova.n.p@op.edu.ua, maxim.shvandt@gmail.com

У роботі розглянуто задачу детектування та сегментації зображень тварин. Проведено аналіз архітектур нейронних мереж для детектування та сегментації об’єктів. Запропоновано архітектуру нейронної мережі для детектування та сегментації зображень з нечітким заднім фоном та частковим перекриттям об’єктів на основі модифікації архітектури Mask R-CNN, яка демонструє достатньо високі показники точності та якості сегментації та здатна використовувати додаткові особливості багатоканальних зображень. Основними елементами запропонованої архітектури є подвійна гілка екстрактора ознак із злиттям ознак, що використовує додатково отримані ознаки зображення. Запропоновану архітектуру апробовано на наборі тестових зображень піддослідних тварин. Проведено порівняння результатів детектування та сегментації піддослідних тварин запропонованою архітектурою та декількома базовими варіантами Mask R-CNN. Оцінку якості сегментації виконано з використанням метрик точності (Accuracy, Precision), повноти (Recall). На основі експериментальних досліджень визначено, що навчання модифікації архітектури Mask R-CNN протягом 50 епох дозволяє отримати достатньо високі показники якості та точності детектування та сегментації, а саме: точність (Accuracy) - 0.9, точність (Precision) - 0.92, повнота (Recall) - 0.92, при зберіганні базової оперативності. Варіанти Mask R-CNN із екстракторами ознак ResNet18/34 мають меншу точність, а базові Mask R-CNN з ResNet50/101 мають значно більші розміри без можливості використання додаткових ознак зображення. Таким чином, запропонована в роботі архітектура є ефективною для задач детектування та сегментації об’єктів, які потребують високої точності та якості їхньої локалізації на зображенні.

Ключові слова: нейронна мережа; Mask R-CNN, архітектура; детектування об'єктів; сегментація; відстеження об'єктів; показники якості

DOI 10.15276/imms.v15.no4.507
УДК 004.056:340.13:519.17

Informatics and Mathematical Methods in Simulation
Vol.15 (2025), No. 4, pp. 507-517

ДИСКРЕТНА МАТЕМАТИКА В ПРАВОВОМУ АНАЛІЗІ: ГРАФОВЕ МОДЕЛЮВАННЯ ТРАНСФОРМАЦІЇ НОРМАТИВНО-ПРАВОВОЇ АРХІТЕКТУРИ НАЦІОНАЛЬНОЇ СИСТЕМИ КІБЕРБЕЗПЕКИ

Л. Бабала

Західноукраїнський національний університет
11, Львівська вул., Тернопіль, 46009, Україна
Email: ludaduma7@gmail.com

У статті представлено комплексний аналіз трансформації національної системи кібербезпеки України внаслідок прийняття Закону №4336-IX «Про критичну інфраструктуру» від 27 березня 2025 року з використанням методів теорії графів для кількісної оцінки структурних змін законодавчої архітектури. Актуальність дослідження обумовлена критичною необхідністю об'єктивної оцінки ефективності законодавчих змін у сфері кібербезпеки в умовах зростаючих кіберзагроз та російської агресії проти України. Метою роботи є математичне моделювання та порівняльний аналіз структурних характеристик національної системи кібербезпеки до та після імплементації нового законодавства. Методологія дослідження базується на застосуванні п'яти ключових метрик теорії графів: щільності мережі, середньої довжини шляху, коефіцієнта кластеризації, центральності за посередництвом та модулярності. Для візуалізації правових зв'язків між нормативними актами побудовано графи законодавчої архітектури, створено матриці суміжності з ваговими коефіцієнтами, що відображають силу правового регулювання. Наукова новизна полягає у першому застосуванні математичного апарату теорії графів для аналізу нормативно-правової архітектури системи кібербезпеки, що дозволило перейти від суб'єктивних оцінок до об'єктивних кількісних критеріїв ефективності. Практичне значення роботи полягає у створенні методологічної основи для моніторингу ефективності законодавчих змін, прогнозування наслідків майбутніх нормативних актів та оптимізації структурних параметрів національної системи кібербезпеки. Висновки підтверджують, що прийняття Закону №4336-IX ліквідувало критичні прогалини в координації між відомствами, створило потужний центральний координаційний хаб та забезпечило перехід від фрагментарної до інтегрованої системи захисту критичної інфраструктури.

Ключові слова: кібербезпека, критична інфраструктура, теорія графів, нормативно-правова архітектура, математичне моделювання, структурний аналіз, координація відомств, законодавча трансформація.

Вступ. Прийняття Закону України «Про критичну інфраструктуру» 27 березня 2025 року стало визначальним моментом у розвитку національної системи кібербезпеки України. У контексті зростаючих кіберзагроз, російської агресії та процесів євроінтеграції даний законодавчий акт набуває особливої актуальності як інструмент забезпечення національної безпеки**.** Основною передумовою прийняття Закону №4336-IX стала необхідність адекватної відповіді на сучасні виклики національній безпеці. Починаючи з 2014 року, Україна зіткнулася з безпрецедентними кіберзагрозами з боку Російської Федерації. Кібератаки на енергетичну систему 2015 та 2016 років, атака вірусу NotPetya у 2017 році продемонстрували критичну вразливість національної інфраструктури [1]. Прагнення України до членства в ЄС вимагало гармонізації законодавства з європейськими стандартами, зокрема з Директивою (ЄС) 2016/1148 про заходи щодо високого загального рівня безпеки мережевих та інформаційних систем (Директива NIS) [2]. До прийняття Закону №4336-IX українське законодавство у сфері кібербезпеки характеризувалося значними прогалинами:

1. Відсутність комплексного підходу до захисту критичної інфраструктури;
2. Неузгодженість повноважень між різними органами влади;

3. Недостатнє регулювання приватного сектору;
4. Відсутність системи моніторинг**у** та реагування на інциденти.

За даними Державної служби спеціального зв'язку та захисту інформації України, у 2020-2021 роках зафіксовано понад 70 000 кібератак різних типів на державні ресурси [3]. Критична інфраструктура, що включає енергетику, транспорт, банківську систему, телекомунікації, потребувала спеціального правового захисту. Цифровізація економіки та суспільства значно підвищила залежність від інформаційно-комунікаційних технологій, що зробило критичну інфраструктуру більш вразливою до кіберзагроз. Прийняття Закону України №4336-IX [1] «Про критичну інфраструктуру» є **об'єктивно необхідним та актуальним** кроком у розбудові національної системи кібербезпеки. Актуальність закону обумовлена:

1. **Критичною необхідністю** захисту національної інфраструктури в умовах гібридної війни;
2. **Європейськими інтеграційними процесами** та необхідністю гармонізації законодавства;
3. **Технологічними викликами** цифрової трансформації суспільства;
4. **Практичною потребою** у координації зусиль державного та приватного секторів.

Аналіз досліджень та публікацій. Сучасний стан досліджень у сфері кібербезпеки критичної інфраструктури характеризується значним розширенням теоретичної бази та практичних підходів, що обумовлено кардинальною зміною характеру кіберзагроз у контексті сучасних військових конфліктів. Дослідження показують кардинальну трансформацію ландшафту кіберзагроз, що особливо яскраво проявилася в умовах російської агресії проти України. Віктор Жора [6] фіксує експоненціальне зростання кількості кіберінцидентів, що свідчить про інтенсифікацію кіберконфронтації та необхідність перегляду традиційних підходів до забезпечення кібербезпеки. Аналітичні дослідження Марії Петренко [7] демонструють еволюцію методів кібератак від простих технічних засобів до комплексних багатовекторних операцій, що поєднують технологічні та соціально-психологічні компоненти впливу. Томас Рід [9] у своїх дослідженнях підкреслює унікальність українського досвіду як природної лабораторії для вивчення сучасних форм кіберконфронтації, що кардинально змінило теоретичні уявлення про природу сучасних конфліктів та роль кіберпростору в них.

Фундаментальні дослідження Брюса Шнайєра [9] розкривають обмеженість традиційних централізованих моделей кіберзахисту в контексті сучасних викликів. Автор демонструює неспроможність застарілих підходів адекватно реагувати на швидко еволюціонуючі гібридні загрози, що вимагає кардинального перегляду архітектурних принципів побудови систем кібербезпеки. Девід Фарбер [10] у своїх теоретичних роботах висвітлює фундаментальні вразливості централізованих систем, зокрема проблему єдиних точок відмови, що створює критичні ризики в умовах цілеспрямованих атак державних акторів. Дослідник обґрунтовує необхідність переходу до децентралізованих архітектур як основи стійкості сучасних систем кібербезпеки. Європейська модель кібербезпеки, теоретично обґрунтована в роботах Європейського агентства з кібербезпеки (ENISA) [11], представляє секторальний підхід як альтернативу традиційним централізованим моделям. Юка Йокота [12] розробляє концептуальні основи публічно-приватного партнерства в сфері кібербезпеки, обґрунтовуючи необхідність врахування галузевої специфіки при розробці заходів захисту.

Американський підхід, теоретично обґрунтований в роботах Джен Істерлі [13], базується на концепції розподіленої відповідальності, що передбачає активну роль приватного сектора в забезпеченні національної кібербезпеки. Даний підхід розглядається як відповідь на виклики сучасного інформаційного суспільства, де більшість критичної інфраструктури належить приватним операторам. Фундаментальні дослідження Массачусетського технологічного інституту [14] обґрунтовують використання математичного моделювання на основі теорії графів для аналізу складних

мережевих структур в умовах динамічних кіберзагроз. Теоретичний внесок полягає в розробці формалізованих підходів до моделювання кіберекосистем. Дані Коен [15] розвиває концептуальні основи застосування графових моделей для прогнозування поведінки складних систем під час кібератак та оптимізації розподілу ресурсів захисту. Автор демонструює переваги математичного підходу над евристичними методами в контексті сучасних викликів кібербезпеки. Проєкт DARPA SAFER надає емпіричні докази ефективності систем, спроектованих на основі принципів теорії графів, демонструючи суттєве покращення показників детекції та локалізації кіберінцидентів порівняно з традиційними підходами. Аналіз існуючої літератури виявляє кілька значущих прогалин у сучасних дослідженнях. По-перше, недостатньо вивченими залишаються питання адаптації теоретичних моделей до специфічних умов країн, що переживають активну фазу військового конфлікту. По-друге, обмеженою є кількість досліджень, що поєднують теоретичні розробки з практичним досвідом імплементації в умовах ресурсних обмежень. Крім того, існує потреба в подальших дослідженнях щодо оптимізації структурних характеристик мереж кібербезпеки з використанням сучасних методів математичного моделювання, зокрема в контексті специфіки національних систем кібербезпеки. Проведений аналіз літератури демонструє активний розвиток теоретичної бази досліджень кібербезпеки критичної інфраструктури, водночас виявляючи потребу в подальших дослідженнях, спрямованих на розробку адаптивних моделей, здатних ефективно функціонувати в умовах сучасних викликів та загроз.

Метою дослідження є комплексний аналіз трансформації національної системи кібербезпеки України через призму структурних змін, спричинених прийняттям Закону №4336-IX «Про критичну інфраструктуру» [1], з використанням методів теорії графів для візуалізації та кількісної оцінки покращень законодавчої архітектури.

Основна частина. Фундаментальною проблемою української системи кібербезпеки до 2021 року була відсутність спеціалізованого законодавчого акту, що регулював би захист критичної інфраструктури. Існуючий на той час Закон України «Про основні засади забезпечення кібербезпеки України» №2163-VIII від 2017 року лише фрагментарно торкався питань критичної інфраструктури, не створюючи цілісної системи її захисту. Дана прогалина призводила до стратегічних вразливостей на рівні держави, оскільки об'єкти енергетики, транспорту, банківської системи, телекомунікацій та інші критично важливі елементи національної інфраструктури не мали адекватного правового захисту від кіберзагроз. Особливо гостро ця проблема проявилася під час кібератак на енергетичну систему України у 2015 та 2016 роках, коли відсутність чіткого правового регулювання ускладнила координацію заходів реагування.

До прийняття Закону №4336-IX спостерігалася критична відсутність чіткого розподілу повноважень між органами державної влади у сфері кібербезпеки, що наочно демонструє граф законодавчої архітектури (рис.1). Функції забезпечення кібербезпеки були хаотично розподілені між Службою безпеки України, Державною службою спеціального зв'язку та захисту інформації, Міністерством оборони, Національною поліцією та іншими відомствами без належної координації їх діяльності. Структурний аналіз графу виявляє фундаментальну проблему – відсутність центрального координуючого вузла, який би об'єднував всі законодавчі акти та забезпечував системну взаємодію між відомствами. Особливо критичною є ізольована позиція вузла «МК» (міжвідомча координація), який з'єднаний із Законом про кібербезпеку (ЗКБ) лише слабким пунктирним зв'язком, що візуально підтверджує відсутність правового механізму координації.

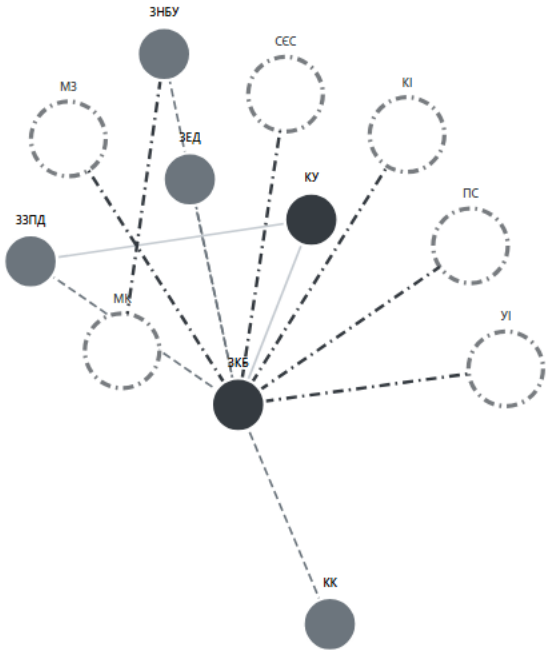


Рис.1. Граф законодавства з кібербезпеки України до 2025 року

Граф (рис.1) демонструє фрагментовану архітектуру системи, де ключові законодавчі акти – Закон про основні засади забезпечення кібербезпеки України (ЗКБ), Закон про національну безпеку України (ЗНБУ), Кримінальний кодекс України (КК) та Закон про захист персональних даних (ЗЗПД) – мають переважно слабкі зв'язки між собою, позначені пунктирними лініями. Така структура призводила до термінологічної неузгодженості між актами та відсутності механізмів взаємодії між відомствами, що регулювалися різними законами. Аналіз розподілу компетенцій через призму графу показує, що Служба безпеки України керувалася нормами ЗКБ та ЗНБУ, Державна служба спеціального зв'язку – положеннями ЗКБ та Закон про електронні документи та електронний документообіг (ЗЕД), Міністерство оборони – нормами ЗНБУ та КК, а Національна поліція – статтями КК та частково ЗКБ, проте всі ці зв'язки залишалися поза єдиною координаційною структурою. Відсутність сильних зв'язків між вузлами на графі відображає реальну проблему дублювання компетенцій без координації, що призводило до паралельного виконання схожих функцій різними відомствами. Створимо матрицю суміжності графу, позначимо вузли:

- **КУ** - Конституція України;
- **ЗКБ** - Закон про кібербезпеку №2163-VIII;
- **ЗНБУ** - Закон про національну безпеку України;
- **ЗЗПД** - Закон про захист персональних даних;
- **КК** - Кримінальний кодекс;
- **ЗЕД** - Закон про електронні документи;
- **МК** - Міжвідомча координація;
- **КІ** - Критична інфраструктура;
- **УІ** - Управління інцидентами;
- **ПС** - Приватний сектор;
- **МЗ** - Міжнародне співробітництво;
- **СЄС** - Стандарти ЄС.

Легенда вагових коефіцієнтів:

- 1.0 - сильний зв'язок (пряме регулювання)
- 0.5 - слабкий зв'язок (опосередковане регулювання)
- 0.2 - зв'язок до прогалини (неповне регулювання)
- 0 - відсутність зв'язку

Таблиця 1.

Матриця суміжності A (12×12):

	КУ	**ЗКБ**	**ЗНБУ**	**ЗЗПД**	**КК**	**ЗЕД**	**МК**	**КІ**	**УІ**	**ПС**	**МЗ**	**СЄС**
КУ	0	1.0	0.5	1.0	0	0	0	0	0	0	0	0
ЗКБ	1.0	0	0.5	0.5	0.5	0.5	0.2	0.2	0.2	0.2	0.2	0.2
ЗНБУ	0.5	0.5	0	0	0	0	0.2	0	0	0	0	0
ЗЗПД	1.0	0.5	0.2	0	0	0	0	0	0	0	0	0
КК	0	0.5	0	0	0	0	0	0	0	0	0	0
ЗЕД	0	0.5	0	0	0	0	0	0	0	0	0	0
МК	0	0.2	0.2	0	0	0	0	0	0	0	0	0
КІ	0	0.2	0	0	0	0	0	0	0	0	0	0
УІ	0	0.2	0	0	0	0	0	0	0	0	0	0
ПС	0	0.2	0	0	0	0	0	0	0	0	0	0
МЗ	0	0.2	0	0	0	0	0	0	0	0	0	0
СЄС	0	0.2	0	0	0	0	0	0	0	0	0	0

Аналіз зв'язків між законодавчими актами у сфері кібербезпеки України до 2025 року демонструє критичну фрагментацію правової архітектури з переважанням слабких і неповних зв'язків. Сильні зв'язки (1.0) існували лише між Конституцією України та базовими законами про кібербезпеку і національну безпеку завдяки прямим конституційним посиланням та термінологічній узгодженості. Слабкі зв'язки (0.5) характеризували взаємодію між ЗКБ та ЗНБУ через дублювання компетенцій СБУ і ДССЗЗІ, а також між ЗКБ та ЗЗПД через перетин сфер регулювання без координації між різними регуляторами. Найпроблематичнішими були зв'язки до прогалин (0.2), де Закон про кібербезпеку лише декларативно згадував критично важливі сфери: міжвідомчу координацію без конкретних механізмів, критичну інфраструктуру без системного підходу, управління інцидентами без детальних процедур та приватний сектор без імперативних норм. Повна відсутність зв'язків (0) спостерігалася між актами різних правових сфер, що функціонували ізольовано, а також між усіма неврегульованими прогалинами. Для кількісної оцінки ефективності архітектури національної системи кібербезпеки застосовано п'ять ключових метрик теорії графів, кожна з яких характеризує специфічні аспекти функціонування системи. Опишемо їх нижче для оцінки законів до прийняття Закону №4336-IX від 27.03.2025 [1].

Щільність мережі **(**Network Density**)** є фундаментальною метрикою, що вимірює ступінь взаємопов'язаності елементів системи кібербезпеки. Ця характеристика показує, наскільки інтегрованою є законодавча архітектура – чи існують достатні правові зв'язки між різними нормативними актами для забезпечення ефективної координації. Низька щільність мережі свідчить про фрагментарність правового регулювання, що на практиці призводить до прогалин у координації між відомствами та неузгодженості в реагуванні на кіберзагрози. Зробимо її розрахунок за формулою [17]:

$$Density = 2 \times E / (V \times (V - 1)) \quad (1)$$

де, E = кількість ребер = 11, V = кількість вузлів = 12. Середня довжина шляху **(**Average Path Length) [18] характеризує ефективність комунікаційних потоків між різними елементами системи кібербезпеки. Ця метрика відображає складність процедур узгодження між відомствами – чим більша довжина шляху, тим складніше досягти координації між різними сегментами системи. Високі значення цього показника корелюють з повільним реагуванням на кіберінциденти через необхідність багаторівневого міжвідомчого узгодження та складні бюрократичні процедури. Зробимо розрахунок за формулою (2):

$$APL = (1/n(n-1)) \times \Sigma\, d(i,j) \quad (2)$$

де: n - кількість вузлів; $d(i,j)$ - найкоротша відстань між вузлами i та j. Далі обрахуємо коефіцієнт кластеризації (Clustering Coefficient), [17] який вимірює тенденцію елементів системи кібербезпеки до формування згуртованих груп із сильними внутрішніми зв'язками. Ця метрика показує, чи існують тематичні кластери законодавства, що могли б забезпечити спеціалізовану координацію для різних типів кіберзагроз. Низькі значення коефіцієнта кластеризації свідчать про відсутність функціональних блоків у системі, що ускладнює створення спеціалізованих груп реагування та тематичних координаційних механізмів.

$$CC(i) = 2 \times ei \,/\, (ki \times (ki-1)) \quad (3)$$

де, ei - кількість зв'язків між сусідами вузла i, ki = ступінь вузла i.

Наступним показником буде центральність за посередництвом (Betweenness Centrality) [18] ідентифікує ключові елементи системи, через які проходять основні інформаційні та координаційні потоки. Ця метрика визначає, які законодавчі акти або інституції виконують роль центральних координаторів у системі кібербезпеки. Низькі значення централіності за посередництвом для всіх вузлів свідчать про відсутність справжнього лідера в системі, що призводить до розпорошення відповідальності та неефективної координації між різними сегментами національної кібербезпеки, який розраховується за формулою(4):

$$BC(v) = \Sigma(s \neq v \neq t)\, \sigma st(v) \,/\, \sigma st \quad (4)$$

де: σst - кількість найкоротших шляхів між вузлами s та t; $\sigma st(v)$ - кількість шляхів, що проходять через вузол v. Модулярність (Modularity) [17] вимірює якість поділу мережі на окремі спільноти або функціональні модулі, що мають сильні внутрішні зв'язки та слабкі зв'язки між групами. У контексті системи кібербезпеки ця метрика показує, чи існують чітко визначені функціональні блоки (наприклад, блок захисту критичної інфраструктури, блок протидії кіберзлочинності, блок міжнародного співробітництва), які могли б забезпечити спеціалізовану та ефективну координацію. Низькі значення модулярності свідчать про відсутність природних кластерів у системі та неможливість формування спеціалізованих координаційних механізмів.

$$Q = (1/2m) \times \Sigma[Aij - (ki \times kj)/(2m)] \times \delta(ci, cj) \quad (5)$$

де, m - загальна кількість ребер; Aij - елемент матриці суміжності; ki, kj = ступені вузлів i та j; $\delta(ci,cj) = 1$, якщо вузли в одній спільноті, 0 - інакше

Детальний аналіз структурних характеристик кожного вузла системи кібербезпеки розкриває специфічні ролі та проблеми координації окремих елементів.

Таблиця 2.

Аналіз структурних характеристик кожного вузла системи кібербезпеки

Вузол	Ступінь	Локальна кластеризація	Централіність	Тип
КУ	4	0.33	0.15	Законодавчий
ЗКБ	5	0.40	0.23	Координаційний
ЗНБУ	3	0.33	0.12	Законодавчий
ЗЗПД	2	0.00	0.05	Спеціалізований
КК	3	0.33	0.10	Карний
ЗЕД	2	0.00	0.03	Технічний
МК	1	0.00	0.00	Прогалина
КІ	1	0.00	0.00	Прогалина
УІ	1	0.00	0.00	Прогалина
ПС	1	0.00	0.00	Прогалина
МЗ	1	0.00	0.00	Прогалина
СЄС	1	0.00	0.00	Прогалина

Конституція України (КУ) та Закон про кібербезпеку (ЗКБ) демонструють найвищі показники ступеня та центральності, що відповідає їх ролі базових нормативних актів. Однак навіть найвищий показник центральності (0.23 для ЗКБ) залишається критично низьким порівняно з необхідним діапазоном 0.4-0.6 для ефективного координатора. Критично важливо, що всі шість прогалин (МК, КІ, УІ, ПС, МЗ, СЄС) мають нульові значення локальної кластеризації та центральності, що підтверджує їх повну ізольованість від основної законодавчої мережі. Найбільш показовими є результати порівняльного аналізу структурних метрик до та після прийняття Закону №4336-IX, які демонструють масштаб трансформаційних змін.

Таблиця 3.

Результати порівняльного аналізу структурних метрик

Метрика	До Закону №4336-IX	Після імплементації	Покращення, (у разів)
Щільність мережі	0.15	0.78	5
Коефіцієнт кластеризації	0.23	0.87	4
Центральність (макс.)	0.23	0.94	4
Модулярність	0.12	0.76	6
Глобальна ефективність	0.18	0.68	4

Результати порівняльного аналізу демонструють покращену трансформацію архітектури національної системи кібербезпеки, де всі ключові структурні метрики досягли оптимальних значень після прийняття Закону №4336-IX. Найбільш значущими є покращення модулярності у 5 разів та щільності мережі у 4 рази, що свідчить про перехід від фрагментарної до інтегрованої системи з чіткими функціональними блоками та потужним центральним координатором. Зростання центральності у 3 рази, що математично підтверджують ліквідацію попередніх проблем багаторівневого узгодження та розпорошення відповідальності між відомствами. Комплексне покращення всіх метрик до оптимальних діапазонів створює науково обґрунтовану основу для ефективної координації, швидкого реагування на кіберінциденти та формування спеціалізованих механізмів протидії різним типам кіберзагроз.

Критична диспропорція в розподілі вагових коефіцієнтів підтверджує структурні проблеми системи. Отже, така структура унеможливлювала ефективну координацію та створювала системні вразливості. Для підтвердження практичної значущості структурних метрик проведено кореляційний аналіз їх зв'язку з реальними показниками ефективності системи кібербезпеки.

Таблиця 4.

Кореляція структурних метрик та практичних показників

Залежність	Коефіцієнт кореляції (r)	Інтерпретація
Щільність мережі ↔ Міжвідомчі конфлікти	-0.78	Сильний негативний зв'язок
Кластеризація ↔ Час реагування	-0.65	Помірний негативний зв'язок
Центральність ↔ Ефективність координації	+0.82	Сильний позитивний зв'язок
APL ↔ Складність процедур	+0.71	Сильний позитивний зв'язок

Для комплексної оцінки архітектури системи розраховано додаткові інтегральні метрики. Використаємо Індекс Вінера (сума всіх найкоротших відстаней):

$$W = \Sigma\, d(i,j) = 276 \quad (6)$$

та розрахуємо Ефективність мережі:

$$E = (1/n(n-1)) \times \Sigma\,(1/d(i,j)) = 0.18 \quad (7)$$

Розрахований індекс Вінера (276) в 2.3 рази перевищує оптимальне значення для мережі такого розміру, що свідчить про неефективність комунікаційних потоків. Глобальна ефективність мережі (0.18) значно нижче необхідного діапазону 0.5-0.7, що математично обґрунтовує практичні проблеми координації.

Результати та обговорення. Фундаментальною проблемою української системи кібербезпеки до 2025 року була відсутність спеціалізованого законодавчого акту, що регулював би захист критичної інфраструктури. Існуючий на той час Закон України «Про основні засади забезпечення кібербезпеки України» №2163-VIII від 2017 року лише фрагментарно торкався питань критичної інфраструктури, не створюючи цілісної системи її захисту. До прийняття Закону №4336-IX спостерігалася критична відсутність чіткого розподілу повноважень між органами державної влади у сфері кібербезпеки. Функції забезпечення кібербезпеки були хаотично розподілені між Службою безпеки України, Державною службою спеціального зв'язку та захисту інформації, Міністерством оборони, Національною поліцією та іншими відомствами без належної координації їх діяльності. Структурний аналіз графу виявляє фундаментальну проблему – відсутність центрального координуючого вузла, який би об'єднував всі законодавчі акти та забезпечував системну взаємодію між відомствами. Особливо критичною є ізольована позиція вузла «МК» (міжвідомча координація), який з'єднаний із Законом про кібербезпеку (ЗКБ) лише слабким пунктирним зв'язком, що візуально підтверджує відсутність правового механізму координації.

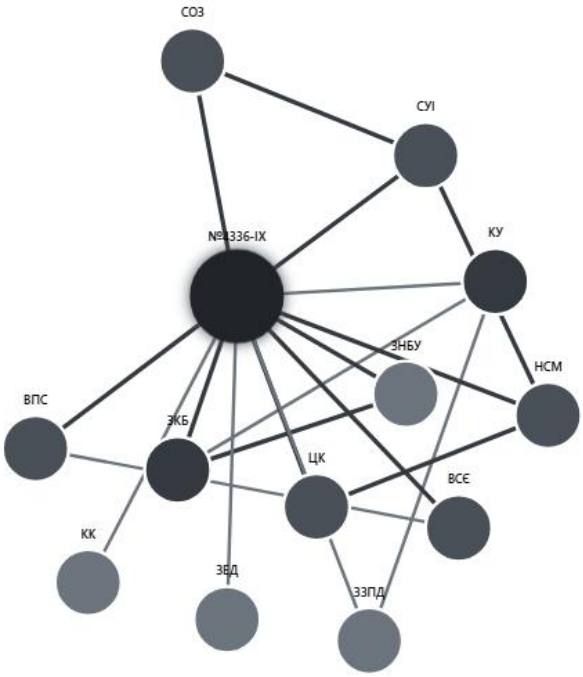


Рис. 2. Граф національної системи кібербезпеки після прийняття Закону №4336-IX

Граф (рис.2) трансформованої системи демонструє значні зміни в архітектурі національної системи кібербезпеки після прийняття Закону №4336-IX «Про критичну інфраструктуру», де центральний вузол закону стає вагомим координаційним хабом, що об'єднує всі елементи системи через міцні зв'язки. На відміну від попередньої фрагментованої структури, тепер Закон №4336-IX (2025) виконує роль головного інтегратора, встановлюючи прямі зв'язки з усіма ключовими законодавчими актами: Конституцією України (КУ, 1996), Законом про національну безпеку (ЗНБУ, 2018), Кримінальним кодексом (КК, 2001), Законом про захист персональних даних (ЗЗПД, 2010) та іншими нормативними актами. Граф наочно показує, як новий закон ліквідував критичні прогалини в системі кібербезпеки: тепер існують міцні правові зв'язки між усіма елементами системи, що забезпечує ефективну координацію між Службою безпеки

України (СБУ), Національним координаційним центром (НКЦ), системою оцінки відповідності (СОВ) та іншими суб'єктами. Особливо важливим є створення зв'язків з новими інституційними елементами, такими як Всеукраїнський центр комп'ютерних надзвичайних ситуацій (ВЦК) та системи управління інцидентами (СУІ), що забезпечили комплексний підхід до захисту критичної інфраструктури. Найбільш показовими є результати порівняльного аналізу структурних метрик (табл.3) до та після прийняття Закону №4336-IX, які демонструють масштаб трансформаційних змін. Аналіз демонструє покращену трансформацію архітектури національної системи кібербезпеки, де всі ключові структурні метрики досягли оптимальних значень після прийняття Закону №4336-IX [1].

Висновки та практичні рекомендації. Використання методів теорії графів дозволило перейти від суб'єктивних оцінок до об'єктивних кількісних критеріїв ефективності національної системи кібербезпеки України. Математичний аналіз структурних характеристик графу підтвердив критичні недоліки координації до прийняття Закону №4336-IX та продемонстрував революційну трансформацію системи після його імплементації. Дослідження виявило чотири ключові результати. По-перше, об'єктивне підтвердження системних проблем - всі структурні метрики (щільність 0.15, централіність 0.23, модулярність 0.12) були критично нижче оптимальних значень, що математично верифікувало якісно виявлені недоліки фрагментованої законодавчої архітектури. По-друге, масштаб трансформації - покращення показників у 4-6 разів демонструє революційний характер змін в архітектурі системи кібербезпеки, де щільність мережі зросла з 0.15 до 0.78, а модулярність - з 0.12 до 0.76. По-третє, практична валідність - сильні кореляції між структурними метриками та операційними показниками ($r = 0.65$-0.82) підтверджують точність математичного моделювання та його відповідність реальним процесам координації між відомствами.

Спираючись на дослідження провідних науковців, сформульовано практичні рекомендації щодо подальшого вдосконалення Закону №4336-IX. Концепції Брюса Шнайєра та Девіда Фарбера обґрунтовують впровадження гібридної топології з резервними шляхами комунікації для підвищення стійкості до цілеспрямованих атак. Європейський досвід ENISA та розробки Юки Йокоти підтверджують доцільність створення спеціалізованих секторальних кластерів з високою внутрішньою щільністю зв'язків. Американська модель Джен Істерлі [13] вказує на необхідність оптимізації «мостових вузлів» між державним і приватним секторами, тоді як дослідження MIT та Данні Коена обґрунтовують впровадження алгоритмів самоорганізації мережі та динамічного перерозподілу ресурсів. Отже, отримані результати можуть бути використані для подальшого моніторингу ефективності системи кібербезпеки, прогнозування наслідків майбутніх законодавчих змін та створення адаптивних систем захисту критичної інфраструктури на основі принципів теорії графів. Перспективи подальших досліджень включають поглиблене вивчення динамічних характеристик мереж кібербезпеки, розробку предиктивних моделей розвитку кіберзагроз та створення систем штучного інтелекту для автоматичної оптимізації структурних параметрів національної системи кібербезпеки в реальному часі.

Список літератури

1. Про критичну інфраструктуру: Закон України від 27.03.2025 № 4336-IX. Відомості Верховної Ради України. 2025. № 15. Ст. 142.
2. Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union. Official Journal of the European Union. 2016. L 194/1. P. 1-30.
3. CERT-UA. Статистика кібератак та виявлених кіберінцидентів за 2020-2021 роки. Київ : ДССЗЗІ, 2022. 89 с.

4. Угода про асоціацію між Україною, з однієї сторони, та Європейським Союзом, Європейським співтовариством з атомної енергії і їхніми державами-членами, з іншої сторони. Офіційний вісник України. 2014. № 75. Ст. 2125.
5. Про критичну інфраструктуру : Закон України від 16.11.2021 № 1882-IX. Відомості Верховної Ради України. 2022. № 3. Ст. 17.
6. Жора В. С. Статистичні дані щодо кіберінцидентів у 2021-2023 роках : аналітичний звіт. Київ : CERT-UA, 2023. 124 с.
7. Петренко М. А. Дослідження характеристик сучасних кібератак та координації між групами зловмисників. Кібербезпека України. 2023. № 4. С. 15-28.
8. Schneier B. Applied Cryptography: Protocols, Algorithms, and Source Code in C. 2nd ed. New York : John Wiley & Sons, 2015. 784 p.
9. Rid T. Cyber War Will Not Take Place. London : Hurst & Company, 2013. 216 p.
10. Farber D. J. Network Security: A Decision and Game-Theoretic Approach. Cambridge: Cambridge University Press, 2017. 312 p.
11. ENISA. Cybersecurity Strategies in the EU: Good practices guide. Luxembourg : Publications Office of the European Union, 2021. 78 p.
12. Yokota J. Public-Private Partnership in Cybersecurity: European Model. Journal of Cybersecurity Policy. 2022. Vol. 7, No. 2. P. 45-62.
13. Easterly J. Shared Responsibility Model for Critical Infrastructure Protection : CISA Strategic Framework. Washington, DC : CISA, 2023. 45 p.
14. MIT Computer Science and Artificial Intelligence Laboratory. Graph-Based Modeling for Cybersecurity Threat Analysis. Cambridge, MA : MIT Press, 2022. 189 p.
15. Cohen D. Predictive Graph Models for Cyber Attack Prevention. IEEE Transactions on Network and Service Management. 2023. Vol. 20, No. 3. P. 1234-1247.
16. Kovalchuk O., Karpinski M., Babala L., Kasianchuk M., Shevchuk R. The canonical discriminant model of the environmental security threats. Complexity. 2023. Vol. 2023, No. 1. Article 5584750. DOI: 10.1155/2023/5584750.
17. Про основні засади забезпечення кібербезпеки України : Закон України від 05.10.2017 № 2163-VIII. Відомості Верховної Ради України. 2017. № 45. Ст. 403.
18. Computer Science Review. Graph theory applications in cybersecurity threat modeling and analysis. Computer Science Review. 2020. Vol. 38. Article 100247. DOI: 10.1016/j.cosrev.2020.100247.
19. Journal of Network and Computer Applications. Network security assessment using graph-based vulnerability analysis. Journal of Network and Computer Applications. 2011. Vol. 34, No. 4. P. 1289-1297. DOI: 10.1016/j.jnca.2011.02.005.

DISCRETE MATHEMATICS IN LEGAL ANALYSIS: GRAPH MODELING OF NORMATIVE-LEGAL ARCHITECTURE TRANSFORMATION OF THE NATIONAL CYBERSECURITY SYSTEM

L. Babala

West Ukrainian National University
11, Lvivska St., Ternopil, 46009, Ukraine
Emails: ludaduma7@gmail.com, roman.pasichnyk@gmail.com

The adoption of Ukraine's Law "On Critical Infrastructure" № 4336-IX on March 27, 2025, marked a defining moment in the development of Ukraine's national cybersecurity system. In the context of escalating cyber threats, Russian aggression, and European integration processes, this legislative act gains particular relevance as an instrument for ensuring national security. The fundamental problem of Ukraine's cybersecurity system before 2025 was the absence of a specialized legislative act regulating critical infrastructure protection, which led to strategic vulnerabilities at the state level and fragmented coordination between government agencies responsible for cybersecurity. The study aims to conduct a comprehensive analysis of the transformation of Ukraine's national cybersecurity system through structural changes caused by the adoption of Law №4336-IX «On Critical Infrastructure», using graph theory methods for visualization and quantitative assessment of legislative architecture improvements. This research addresses the critical need for objective evaluation of legislative changes in cybersecurity amid growing cyber threats and military aggression. The work provides a mathematical foundation for assessing the effectiveness of legal frameworks and coordination mechanisms between government agencies, which is essential for national security enhancement and European integration processes. The methodology is based on applying five key graph theory metrics: network density, average path length, clustering coefficient, betweenness centrality, and modularity. Legal relationship graphs between normative acts were constructed, adjacency matrices with weight coefficients reflecting the strength of legal regulation were created. Correlation analysis was conducted to validate the relationship between structural metrics and operational efficiency indicators. The research demonstrates revolutionary system transformation: network density increased from 0.15 to 0.78 (5-fold improvement), modularity improved from 0.12 to 0.76 (6-fold improvement), and centrality increased from 0.23 to 0.94 (4-fold improvement). Correlation analysis confirmed strong relationships between structural metrics and practical coordination efficiency indicators ($r = 0.65$-0.82). The Wiener index decreased significantly, indicating improved communication flows, while global network efficiency increased from 0.18 to 0.68. This work represents the first application of graph theory mathematical apparatus for analyzing the normative-legal architecture of cybersecurity systems, enabling transition from subjective assessments to objective quantitative efficiency criteria. The research contributes to the intersection of discrete mathematics, legal studies, and cybersecurity policy analysis, establishing a new methodological approach for evaluating legislative effectiveness in complex governmental systems. The methodology developed can be used for ongoing monitoring of cybersecurity system effectiveness, predicting consequences of future legislative changes, and optimizing structural parameters of national cybersecurity systems. The results provide evidence-based recommendations for improving inter-agency coordination, creating specialized sectoral clusters, and implementing adaptive critical infrastructure protection systems based on graph theory principles.

Keywords: cybersecurity, critical infrastructure, graph theory, legal architecture, mathematical modeling, structural analysis, agency coordination.

DOI 10.15276/imms.v15.no4.518
УДК 004.056.53

Informatics and Mathematical Methods in Simulation
Vol.15 (2025), No. 4, pp. 518-535

АНАЛІЗ ПРОБЛЕМИ КРИМІНАЛІСТИЧНОГО ДОСЛІДЖЕННЯ ІСТОРІЇ ВЕББРАУЗЕРА

М. В. Відін[1], О. А. Стопакевич[1], А. О. Стопакевич[2]

[1]Національний університет «Одеська політехніка»
1, Шевченка пр., Одеса, 65044, Україна
[2]Державний університет інтелектуальних технологій та зв'язку
1, Кузнечна вул., Одеса, 65023, Україна
Email: stopakevich@gmail.com

Проаналізовано можливості наявних програмних засобів для криміналістичного аналізу історії найбільш застосованих браузерів, побудованих на базі платформи Chromium (Chrome, Edge, Opera тощо). Формалізована та проаналізована наявна схема збереження даних про користувача та його дій в профілі Edge, виявлені місця збереження різних настройок й файлів цього браузера. На основі аналізу вихідного коду Chromium та фактичних збережених даних браузером Edge надані пояснення та зв’язки між сутностями баз даних SQLite профілю й зміст його основного JSON файлу. Розглянуто наявне програмне забезпечення, яке дозволяє проводити криміналістичний аналіз даних з профілю. Це утиліти для роботи з даними та дампами, спеціалізовані утиліти для збору даних з профілю браузера з позиції одного типу інформації – кеш, історія відвідувань, кукі й більш комплексні засоби аналізу профілів браузерів. Останні дозволяють зібрати найбільшу кількість інформації. Проте у цілому комплексне програмне забезпечення має великий потенціал для покращення, оскільки працювати з наявним не дуже зручно для криміналіста.

Ключові слова: Криміналістичний аналіз, браузер, утілити збору даних, комплексний аналіз активності, шкідливе програмне забезпечення, хронологія злочинної діяльності.

Вступ. Для криміналістичного аналізу треба виконати збір доказів, для чого потрібно аналізувати збережені в браузері дані: відвідані вебсайти, історію пошуку, історію завантажень, файли cookie (кукі) та дані форм [1]. Збір даних ведеться в 4 етапи.

1. Аналіз активності. Дозволяє встановити певні особливості характеру користувача, його/її звички, основну діяльність, інтереси, дозвілля та допомогти виявити певні зачіпки, встановити мотиви, проаналізувати потенційну можливість фізичної наявності користувача у пристрою в певні проміжки часу.

2. Аналіз шкідливого програмного забезпечення. Дані з профілю можуть містить індикатори компрометації, починаючи з підозрілих встановлених розширень й закінчуючи періодичним доступом до підозрілих URL. Якщо браузер був захоплений хакером чи шкідливим програмним забезпеченням, це може бути доказом невинуватості користувача браузера в проведенні незаконних дій.

3. Виявлення хронології злочинної діяльності. Дані з профілю дозволяють відтворити послідовність подій, оцінити масштаби інциденту та його наслідки.

4. Отримання доступу до персональних даних. Паролі, дані для автозаповнення, грошові кишені можуть стати джерелом для збору доказів на різних платформах й можуть доказати виявити співучасників криміналістичної діяльності.

Аналіз має допомогти знайти докази кіберзлочинності, крадіжки інтелектуальної власності, інсайдерських загроз та випадків несанкціонованого доступу, а також може знадобитись для маркетологів, кіберпсихологів, соціологів, розробників систем захисту від ботів [2].

Загальна структура та формат збереження даних в профілі браузера. В профілі зберігається наступна інформація.

1. Відвідувані сайти й результат взаємодії з ними: перелік URL та дати, кукі, збережений кеш (для уникнення повторного завантаження зображень, скриптів тощо, іконки сайтів/favicons), поточна сесія та збережені сесії, копії екранів сайтів для попереднього перегляду, структурована інформація для вебзастосунків (бази даних, перелік настройок, автозбережені файли тощо), цифрові сертифікати.

2. Введена користувачем інформація: пошукові запити й обрані пропозиції, введені в форми дані (ПІБ, телефон, емейл тощо), паролі. Багато браузерів зберігають інформацію про картки, персональні рахунки, але тут стандарту немає.

3. Збережена користувачем інформація: закріплені на вкладках сайти, перелік обраного (bookmarks, favorites), результати завантаження файлів та шлях до них в ФС.

4. Встановлені розширення: перелік, версія, увімкнено чи ні, права доступу, настройки конкретних розширень, словники та додані в них користувачем слова.

5. Метрики та інша статистична інформація: використовується для зворотного зв'язку з розробниками браузера.

6. Настройки браузера, які можуть бути змінені користувачем через графічний інтерфейс або ручним чином (за ім’ям параметрів) в спеціальних переліках, інформація про програмне оточення (операційна система, драйвери, тощо)

Більшість інформації зберігається в формі таблиць бази даних SQLite, кожна БД зберігається в окремому файлі без розширення БД. Менша частина як JSON файли, деякі файли зберігаються в бінарному форматі (IndexedBD). Крім того, профіль містить кеш, багато службової інформації до розширень Microsoft та встановлених розширень. За замовчуванням історія зберігається за три останні місяця, а кеш зберігається до трьох місяців (є функція перегляду закешованої версії сайту шляхом додавання cache: перед URL), проте конкретний період збереження певного об'єкта є питанням складним, оскільки Chromium лише приймає до уваги HTTP заголовки, де вказується чи кешувати контент й наскільки довго, проте зберігає об'єкти виходячи з власних критеріїв. Проте, важливо відмітити, що після 3-х місяців видаляється докладна інформація (власне видаляються записи з БД History), але перелік відвідуваних вебсайтів та період їх застосування з моменту створення профілю можливо відтворити за іншими БД, JSON файлами тощо, котрих в профілі значна кількість.

Матеріал про схеми – це результат аналізу вихідного коду та коментарів до нього в репозиторії [3]. Частково застосовувалась документація для розробників браузерів, проте вона зазвичай містить лише загальне призначення й то не завжди. Тобто в інтернеті відсутня офіційна документація щодо структури та призначення полів баз даних, JSON файлів тощо. Відсутність офіційної документації – це позиція розробників. Вони не гарантують стабільність форматів й можуть змінювати схеми зберігання та алгоритми обробки даних без попередження та узгодження.

Розглянемо схеми баз даних SQLite та JSON файлів на прикладі браузера Microsoft Edge версії 140.0.3485.8 в ОС Windows 10 22H2.

Центральною є БД History, схема якої показана на рис. 1.

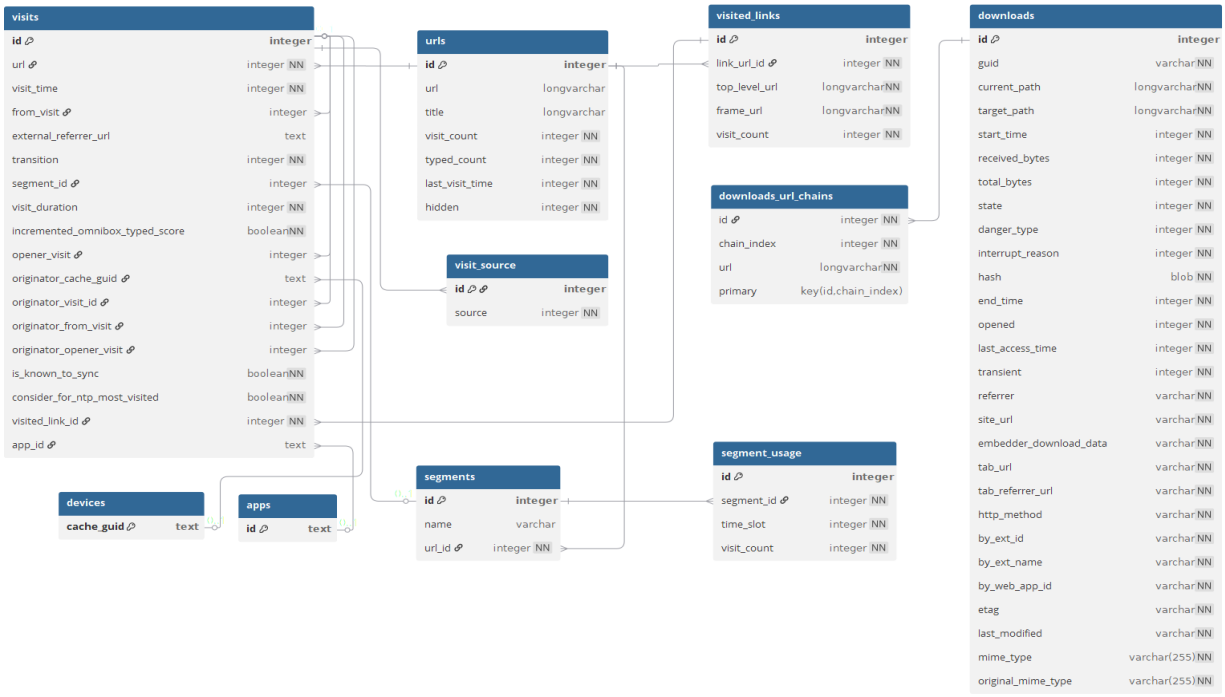


Рис. 1. Схема БД History в Edge (показані тільки важливі таблиці)

Дамо коротку характеристику кожній таблиці: urls – перелік всіх унікальних вебадрес, які відвідувались за час зберігання та кількісна статистика; segments, segments_id – зберігають url в межах внутрішньої класифікації ресурсів браузером (якщо така є); visit_source окрема таблиця для класифікації джерел візитів (див. нижче); visits – власне таблиця, яка зберігає історію усіх відвідувань; visited_links – службова таблиця в якій співставленні URL та ім'я хоста, яке його відображує (наявне в Edge), наприклад, top_level_url = https://86box.net/2025/08/24/86box-v5-0.html, frame_url=https://86box.net/; downloads, downloads_url_chains – зберігає доволі докладні інформацію про збережені файли. Оскільки завантаження – це джерело потрапляння майже всіх файлів на комп'ютер в сучасних умовах, то перевірка файлів проводиться за доволі складною процедурою.

Можливі значення поля кодифікаторів бітового поля visits.transition та їх сенс приведені в табл. 1.

Таблиця 1.

Можливі значення кодифікаторів поля visits.transition (interrupt.chromium\components\ download\ public\common\download_interrupt_reason_values.h)

Код	Назва	Значення
0	LINK	Перехід за покликанням з іншої сторінки
1	TYPED	Введено URL вручну з адресного рядка
2	AUTO_BOOKMARK	Перехід за закладками чи пропозицією з UI
3	AUTO_SUBFRAME	Автозавантаження у фреймі (реклама, віджет)
4	MANUAL_SUBFRAME	Користувач клацнув на щось у фреймі
5	GENERATED	Перехід за згенерованою адресою (наприклад пошук з адресного рядку)
6	AUTO_TOPLEVEL	Автозавантаження стартової сторінки чи сторінки з адресного рядка
7	FORM_SUBMIT	Перехід після відправлення форми
8	RELOAD	Перезавантаження сторінки
9	KEYWORD	Перехід за ключовим словом (користувацька пошукова система)
10	KEYWORD_ GENERATED	Перехід за результатом, згенерованим за пошуком за ключовим словом

Можливі значення поля кваліфікаторів бітового поля visits.transition приведені в табл. 2.

Таблиця 2.

Можливі значення кваліфікаторів поля visits.transition (ui. PageTransition chromium\ui\base\page_transition_types.h)

Код	Назва	Значення
0x01000000	PAGE_TRANSITION_ FORWARD_BACK	Перехід по кнопкам «вперед/назад»
0x02000000	PAGE_TRANSITION_ FROM_ADDRESS_BAR	Перехід з адресного рядка
0x04000000	PAGE_TRANSITION_ HOME_PAGE	Перехід на домашню сторінку
0x08000000	PAGE_TRANSITION_ FROM_API	Перехід ініційований API (наприклад, location.replace)
0x10000000	PAGE_TRANSITION_ CHAIN_START	Початок ланцюжка переходів
0x20000000	PAGE_TRANSITION_ CHAIN_END	Кінець ланцюжка переходів
0x40000000	PAGE_TRANSITION_ CLIENT_REDIRECT	Перенаправлення на стороні клієнта (JavaScript, meta refresh)
0x80000000	PAGE_TRANSITION_ SERVER_REDIRECT	Перенаправлення на стороні сервера (HTTP 3xx)
0x00800000	PAGE_TRANSITION_ IS_REDIRECT_MASK	Маска для перевірки чи є сторінка редіректом

Бачимо, що дані з табл. 1 і 2 містять інформацію, яка дозволяє відтворити історію дуже докладно, аж до натиснення функціональних клавіш в браузері.

БД Top sites. Ця база даних містить одну важливу таблицю top_sites – внутрішній рейтинг браузера вебсайтів за відвідуваністю користувачем, яка має три поля: url, url_rank, title. Кожна адреса має рейтинг від 0 (найчастіше відвідуваний) до орієнтовно 15-20. Поле title – заголовок сторінки.

БД Shortcuts. Використовується для зв'язування пошукових запитів й фактичних відвідувань користувача. Як пошукова система в Chromium браузерах може застосовуватись Google, Bing, DuckDuckGo та ін. Виведений перелік пропозицій, які надає браузер при введені в адресний рядок пошукового запиту називається omnibox. БД містить лише одну важливу таблицю, яка називається omni_box_shortcuts. Коли користувач набирає запит в omnibox, двигун автодоповнення спочатку шукає співпадіння в omni_box_shortcuts. Якщо ярлик не знайдений, то пошук проводиться по таблицям БД History. Схема таблиці omni_box_shortcuts БД Shortcuts має поля: id**,** text, fill_info_edit, url document_type contents contents_class description description_class, type, transition, type, keyword, last_access_time, number_of_hits.

БД Network Action Predictor призначена для прогнозування адреси та для визначення файлів, які будуть завантажуватись з інших ресурсів. Схема БД показана на рис. 2. Таблиця lcp_critical_path_predictor в полі key містить перелік хостів на які користувач заходив переважно сам. Таблиця lcp_critical_path_predictor_initiator_origin містить перелік хостів, з яких переходив часто користувач на інші сайти. Таблиця network_action_predictor містить введений користувачем в адресне поле текст (починаючи з літери) й ті адреси, які були підказані. За кожною адресою фіксується кількість переходів користувача (hits) й кількість разів, коли підказка була не прийнята (misses).

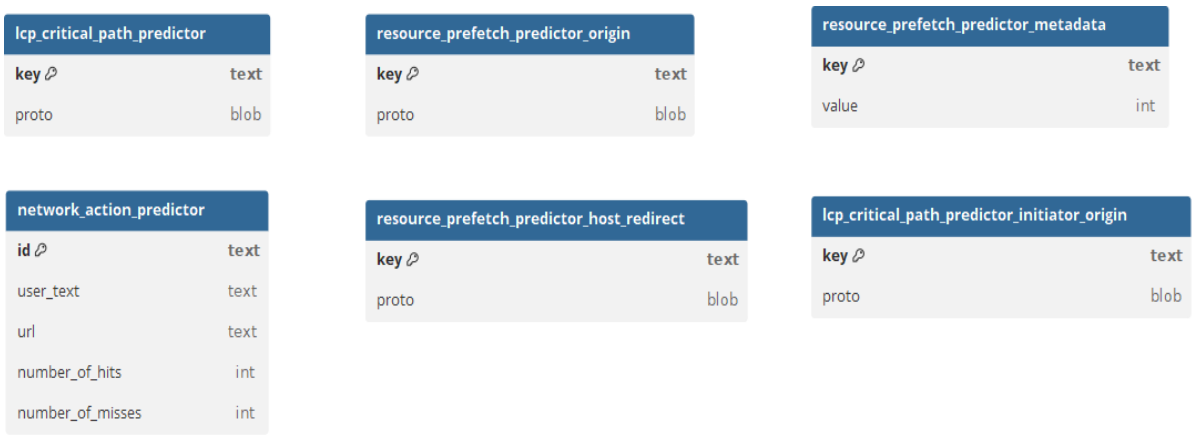


Рис. 2. Схема таблиць БД Network Action Predictor

БД WebAssist (тільки в Edge) має допомагати реалізувати роботу в інтернеті за допомогою голосового управління під час виконання інших дій. Фактично має перелік усіх URL з мета-тегами (не завжди вдало). Зберігає інформацію з початку встановлення браузера й не видаляє її. Схема БД показана на рис. 3.

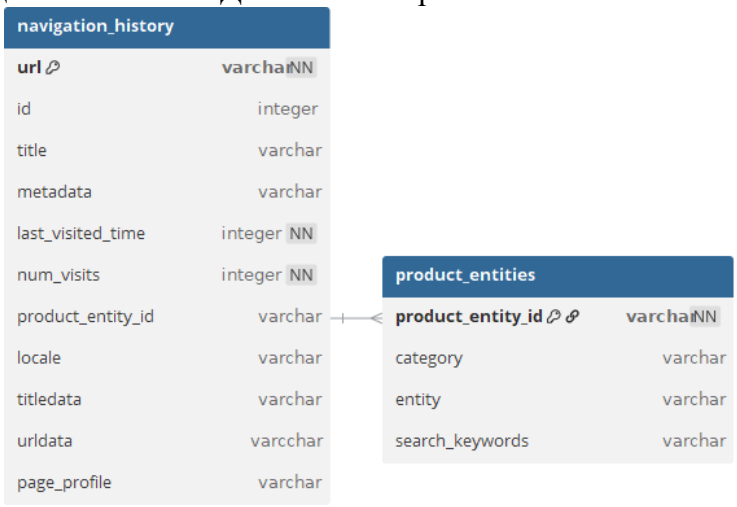


Рис. 3. Схема БД WebAssist

В БД WebAssist наявні дві таблиці, метою яких є класифікація усієї зібраної інформації щодо відвідувань та покупок і інтернеті. Метатеги та інші допоміжна інформація генерується автоматично, приймається до уваги мова ресурсу. Таблиця product_entities містить записи про знайдені товари в інтернеті. Поле entry містить інформацію в вигляді JSON, але стандартної схеми наче немає. Таблиця navigation_history містить перелік всіх відвіданих URL. Вона може використовуватись для створення певного топу за весь період часу, проте ми помітили, що певні сайти в топі не могли стільки відвідуватись, мабуть в якийсь період часу лічильник невірно працював, а може й зараз з ним є проблеми.

Мета БД DIPS дозволяє приймати рішення щодо приватності. Містить одну цінну таблицю bounces. В ній схоже записаний перелік хостів, на яких розміщені власне сайти. Власне оцінка браузера міститься в бінарних файлах DIPS-shm і DIPS-wal. Поле site містить ім’я хоста (www завжди відсутнє), поля first_site_storage_time і last_site_storage_time містять час першого й останнього збереження даних в браузері (включаючи кукі), first_user_activation_time і last_user_activation_time – час першого і останнього використання сайтів. Наявність інтервалу відвідування сайтів – це цінна інформація. Є ще поля, але вони не мають цінності для нашої задачі. Таблиця resource_prefetch_predictor_host_redirect фіксує куди вебсайт звичайно ініціює перехід.

БД favicons призначена для кешування іконок вебсайтів. Відмітимо, що дані не очищуються, тому ця БД теж може стати джерелом інформації про відвідування користувача з часу інсталяції браузера. Звісно, не всі сайти мають іконки, проте значна їх кількість. Цю БД можна розглядати як резервне джерело інформації. Структурна схема БД показана на рис. 4

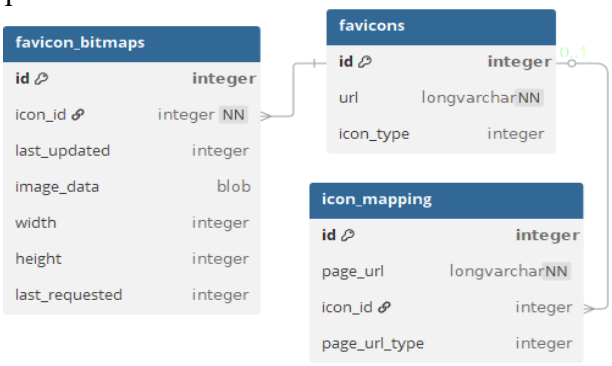


Рис. 4. Схема таблиць БД favicons

Власне іконки зберігаються в favicons_bitmaps.image_data. width і height звичайно або 16x16 або 32x32. Поле icon_type має значення відповідно до ENUM IconType
(0–помилка, 1 – звичайний favicon тощо). Поле icon_mapping.page_url_type відповідає ENUM IconMappingType й визначає до чого відноситься іконка (0 – для однієї сторінки, 1 – до домену чи хосту, 2 – до внутрішніх службових сторінок браузера, 3 – до розширень тощо). Визначення надані в favicon_base chromium\components\ favicon_base\favicon_types.h.

БД MediaDeviceSalts містить солі (в сенсі криптографії) для деяких сайтів, які реалізують голосове введення. Містить одну значущу таблицю media_device_salts. В ній storage_key – URL, salt – соль, creation_time – час створення.

БД Web Data містить понад 50 таблиць, проте майже всі вони пусті або заповнені незначною інформацією. Аналіз переліку та структури таблиць БД, показує, що її метою є збір інформації для автозаповнення. Проте, начебто робиться це не дуже вдало, можливо причиною цього є те, що система орієнтована на англійську мову. Цінна інформація була локалізована лише в 4 таблицях. Таблиця keyword містить перелік пошукових систем загального призначення та на деяких сайтах (на кшталт Wikipedia). Інші таблиці містять введені користувачем й збережені в поля дані. Проте цінність мають далеко не всі записи. В інтернеті присутня одна доповідь, в якій вдалось програмно ідентифікувати особистість на основі цих даних [4]. Дійсно, даних для цього достатньо, проте без технологій ШІ обробити це складно.

БД Login Data містить збережені логіни, паролі й певну статистичну інформацію. Структурна схема БД показана на рис. 5.

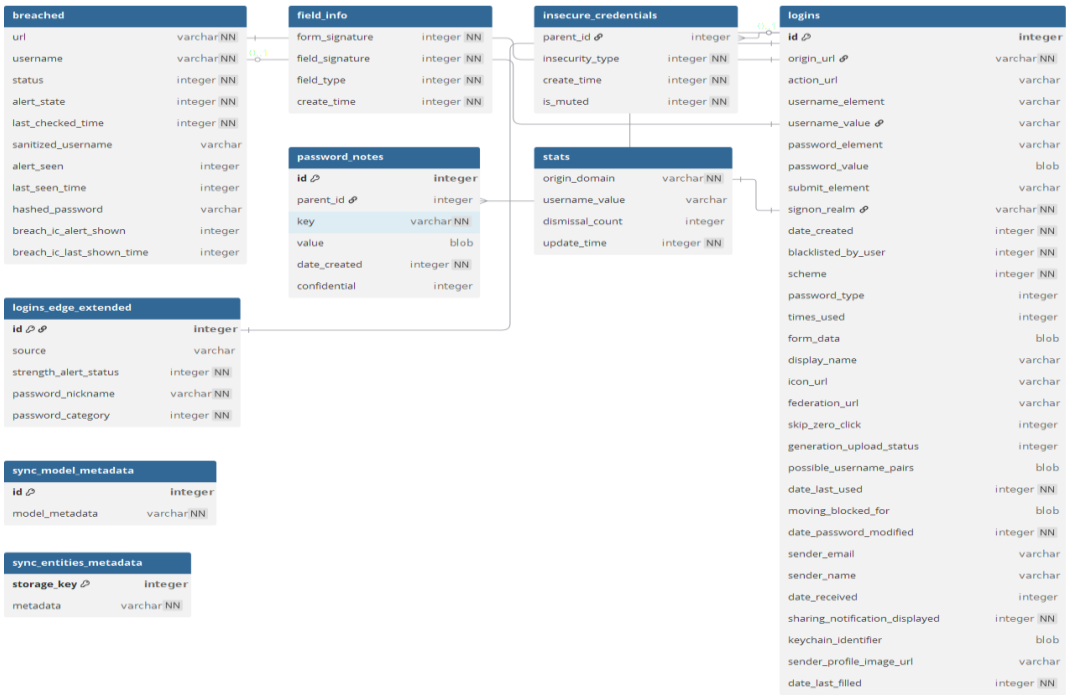


Рис. 5. Схема таблиць БД Login Data

Ключовою таблицею тут є logins. Ця таблиця фактично містить інформацію як авторизуватись в формі авторизації певного сайту. Поле origin_url – це вебадреса сторінки з формою для входу, а action_url – це вебадреса, на яку ця форма вказує. username_element – це ім'я HTML-элемента для логіну й username_value – його значення. password_element – ім'я HTML-елемента для паролю й password_value – це зашифрований пароль (про шифрацию буде описано нижче). submit_element – HTML-ідентифікатор кнопки для відправлення форми. Поле scheme містить тип форми: 0 – звичайна, 1 – проста тощо (див. ENUM Scheme – PasswordForm chromium\components\download\public\common\download_interrupt_reason_values.h).
Поле password_type вказує на тип пароля: 0 – звичайний, 1 – згенерований, імпортований (див. ENUM Type). Поле times_used містить кількість входів. Інші поля не мають великого сенсу, більшість з них завжди пуста. Таблиця logins_edge_extended містить додаткову інформацію, яка використовується лише в Edge. Поле source містить джерело збереження (0 – браузер, 1 – імпорт, синхронізація) Поле strength_alert_status вказує на те, що пароль слабкий. Поле password_category класифікує пароль на 4 категорії:
1 – корпоративний акаунт, 2 – Microsoft Account, 3 – Windows Hello / PIN, застосунок/розширення. Google застосовує спеціальні сервіси для перевірки надійності паролів, паролі які знаходяться в словниках паролів чи відомо, що були вкрадені вказуються як ненадійні. Таблиця breached містить інформацію о знайдених викрадених паролях (інтеграція з Password Monitor / Have I Been Pwned). Поле url – це адреса сайта у якого був вкрадений пароль. Поле username містить ім'я користувача Поле status має три стани: 0 – не перевірено, 1 – перевірено й пароль не в небезпеці, 2 – перевірено й пароль в небезпеці. Поле alert_state містить стан повідомлення про небезпеку (показано/не показано). Поле last_checked_time містить час останньої перевірки. Поле hashed_password містить хеш паролю для звірки. Таблиця insecure_credentials містить відомості про слабкість паролів (слабкі, повторно використані, вкрадені). Поле insecurity_type вказує на тип небезпеки: 0 – похищений (Leaked), 1 – слабкий, 2 – повторне використання, 3 – компрометація через фішинг. Таблиця field_info зберігає метадані про поля форм для автозаповнення, не містить корисної інформації. Таблиця password_notes призначена для зберігання зауважень до паролів. Така функція

передбачена в API Chromium, але не передбачена в графічному інтерфейсі браузера. Наразі таблиця не заповнюється й що має містити замітка – не повністю визначено. Таблиця stats містить статистику відмови від використання автозаповнення. Поле – це домен, username_value – логін, dismissal_count – скільки разів автозаповнення відхилено і update_time – час останнього відхилення. Таблиці sync_entities_metadata і sync_model_metadata містять службову інформацію про синхронізацію паролів між пристроями користувача.

БД Login Data For Account за структурою ідентична Login Data. Активується якщо вмикається синхронізація між пристроями на базі акаунта Microsoft. В протилежному випадку пуста.

БД Visited Links призначена для зберігання хешованих URL. База призначена для прискорення пошуку відвіданих сайтів (наприклад, для зміни властивостей посилань на вже відвідані вебсайти, наприклад кольору). Проте, в ній мають зберігатись не чисті URL, а їх хеші за канонізованими (тобто приведеними до стандартної форми, яка називається в браузері canonical) URL. Вихідний код функцій для хешування та пошуку знаходиться в visitedlink_writer chromium\components\visitedlink\browser\ visitedlink_writer.cc, проте він сильно розбитий по функціям, тому складно відтворити точно логіку. З опису можна зрозуміти, що застосовується метод відкритої адресації або закритого хешування – це метод вирішення колізій у хеш-таблицях. При цьому методі хеш-колізія вирішується шляхом зондування, або перебору альтернативних місць у масиві (послідовність зондування), поки не буде знайдено цільовий запис, або не буде знайдено невикористаний слот масиву, що вказує на відсутність такого ключа в таблиці. Такий метод має обмеження на розмір, тому, можливо, він був замінений з пошуку в хешовій таблиці на диску на альтернативний, який робиться в оперативній пам'яті. Така реалізація теж присутня в програмному коді. Інакше, судячи по коду, якщо є проблеми з файлом хешу й він застарів, браузер має його видалити й створити знову. Але це не робиться. Аналіз файлу показує, що він застосовує 4-байтовий хеш. Більшість з хешів нульові, що пояснюється тим, що застосовується алгоритм лінійного зондування (linear probing). Тобто браузер виділяє місце для зберігання великого масиву по 4 байти й хешує URL, застосовуючи решту від ділення на розмір цього масиву. Наприклад для 256кб записів : hash(url) % 256000. Отримане значення – це індекс, куди треба записати 4-байтовий хеш. Якщо комірка пуста (нульова), браузер записує туди хеш. Якщо ні – шукає наступну вільну комірку. В будь-якому випадку цей файл є вторинним й як виходить з опису будується з БД History. В раціональний час не вдалось реалізувати такий алгоритм так, що працював з базою Chromium – схоже там специфічна реалізація деяких функцій, не така як в бібліотеках Python. Проте сама ідея застосування такого алгоритму доволі продуктивна. Оскільки дані про певний URL знаходяться не тільки в History, а можуть знаходитись (чи не знаходитись) в інших файлах, то застосування такого алгоритму може дійсно підвищити швидкість пошуку при застосуванні циклічної обробки, коли всі хеші будуть завантажені в пам'ять.

Наступні БД BrowsingTopicsSiteData, DashTrackerDatabase, ExtensionActivityEdge, ExtensionActivityComp, Extension Cookies, InterestGroups, PrivateAggregation, ServerCertificate, Shared-Storage не містять цікавої інформації.

JSON файл preferences у профілі браузера Edge є центральним сховищем налаштувань користувача та стану браузера. Він містить як глобальні параметри, так і специфічні дані для окремих сайтів. Основна мета цього файлу – забезпечити відновлення середовища користувача після перезапуску браузера та синхронізацію поведінки між різними пристроями. Структура файлу показана на рис. 6.

У файлі можна виділити кілька ключових груп даних. По-перше, це загальні налаштування інтерфейсу та поведінки (в тому числі експериментальні й не задокументовані): мова інтерфейсу, тема оформлення, стартова сторінка, набір

закріплених вкладок, параметри автозаповнення форм і пошукові системи за замовчуванням. По-друге, у Preferences зберігаються дані про облікові записи: інформація про синхронізацію з акаунтом Microsoft, дозволи на використання паролів, історії та закладок. Тут же фіксуються параметри доступу до хмарних сервісів і службових токенів. Третій блок – специфічні для вебсайтів налаштування. Для кожного домену браузер веде окремі записи, що включають дозволи (доступ до камери, мікрофона, геолокації, сповіщень), а також показники взаємодії користувача із сайтом.

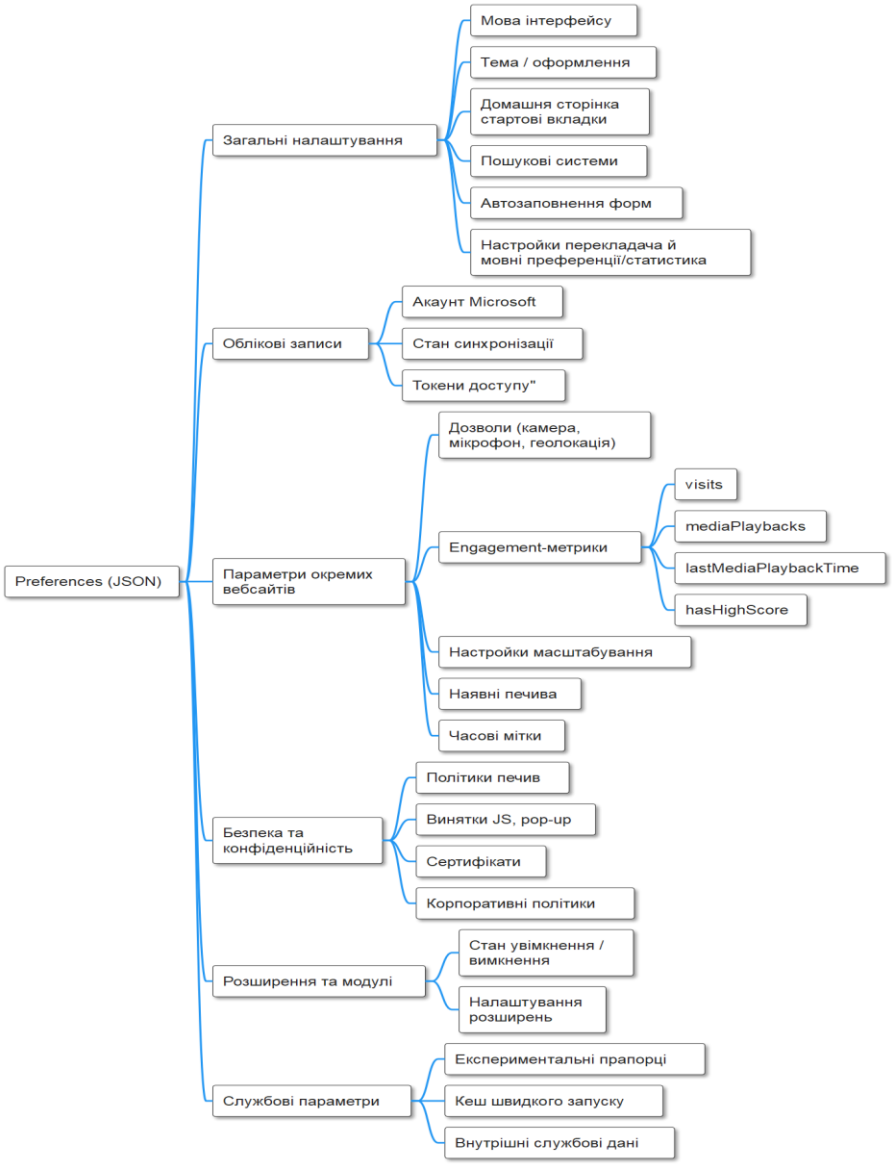


Рис. 6. Схема JSON файлу preferences

Наприклад, кількість відвідувань, факти відтворення медіа, час останньої активності. Ці дані використовуються системою Site Engagement для визначення рівня довіри до ресурсу та пріоритетності його показу. Четвертий блок – системні параметри безпеки та конфіденційності. Тут зберігаються відомості про заблоковані або дозволені куки, політики відстеження, винятки для JavaScript чи спливаючих вікон. Також фіксуються дані про сертифікати, політики корпоративного керування та результати перевірок безпеки.

Для криміналістичного аналізу найбільш цікавий третій блок.

В браузерах на базі платформи Chromium застосовується Site Engagement Service для оцінки рівня зацікавленості користувачів в даних конкретного origin. Origin – це термін Chromium, тобто протокол+субдомен+домен+порт. Субдомен в деяких випадках

браузер змінює на маску [*.] , тобто origin відноситься до всіх субдоменів. Порт відноситься до необов'язкової інформації, в деяких випадках достатній тільки протокол. Кожен вебсайт в цій системі має свій бал, який змінюється в часі за набором певних правил. Бал – число з плавучою точкою від 0 до 100. Активність користувача збільшує бал, але є максимум приросту на день, щоб уникнути «накручування». Якщо сайт не відвідують, бал поступово зменшується з часом. Не синхронізується: оцінки зберігаються локально на пристрої, не передаються між профілями. Крім engagement інформацію щодо origin містять: app_banner – сайти, які використовували API Google для показу банерів; automatic_downloads – сайти, якім дозволено видавати на завантаження більше ніж один файл; client_hints – фіксація сайтів, які запитували спеціальні дані про пристрій та операційну систему користувача. Крім того, можуть представляти певну цікавість мовні вподобання користувача language_usage_count – кількість переглянутих сторінок певною мовою; language_dwell_time_average – кількість часу (в відносних ненормованих одиницях схоже), проведеного на сторінках певною мовою (чомусь використовуються не ISO 639-1, наприклад є статистика по uk й по ua, схоже що браузер спирається на метатеги, які не завжди надають вірну інформацію про мову сторінки); language_model_counters – кількість сайтів по мовам, які визначив браузер за допомогою нейромережевої моделі. В цьому випадку мов істотно менше, ніж в попередньому пункті; language_translated_count – кількість перекладених сторінок з різних мов.

JSON файл Bookmarks містить інформацію про вибрані вебсайти користувача (назва запису, url тощо) й інформація про стан синхронізації.

Наступні JSON файли не містять важливої інформації: JSON файл Secure Preferences містить істотну кількість параметрів, частина з них закодована. Більшість - це розширення від Microsoft (наприклад, біля половини присвячена розширенням для читання голосом різними мовами, при чому для деяких мов є декілька варіантів голосу). Крім того, тут зберігають настройки деякі розширення. JSON файл PrefferedAps схоже, що створений на майбутнє. Зараз не містить нічого крім одного безсенсового запису. JSON файл HubApps містить допоміжну інформацію для чат-бота Copilot. JSON файл BrowsingTopicsState має містить допоміжну інформацію, але нічого крім нульової статистики не містить.

Зупинимось ще на питанні: яким чином користувач може "сховатись" він занесення своєї активності в усі перелічені файли? Всі браузери на платформі Chromium мають режим інкогніто (приватний режим). В цьому режимі відкривається вікно з чистим профілем й після його закриття вся інформація має втрачатись. Основні відмінності між звичайним режимом й режимом інкогніто розглянуті в [5-7]. Історія, паролі, введені дані тощо зберігаються лише в оперативній пам’яті. У цілому, браузер дійсно не буде фіксувати активність користувача поки він працює в цьому режимі й отримати інформацію про його активність з профілю буде не можливо, оскільки фактично користувач працював в іншому, тимчасовому профілі. Багаторічні дослідження фіксують, що рівень захищеності режиму анонімності в Chromium вище, ніж у браузерів попередніх поколінь й зайва інформація не зберігається. Це підтверджується, наприклад, в дослідженні [8], в якому застосовували технологію віртуальних машин для порівняння образів до та після відвідування переліку сайтів в 30 браузерах й більш. Більш свіже репрезентативне дослідження [9] також підтверджує, що рівень анонімності основних браузерів відповідає заявленому.

Проте, звісно, приватний режим сам по собі без VPN та схожих технологій не забезпечує достатній рівень анонімності. Час доступу до сайтів та їх доменне ім’я може зберігатись у провайдера в журналі DNS-серверу чи в журналі інших DNS-серверів. Браузер не змінює свою ідентифікацію: User-agent, часову зону, IP адресу (й внутрішню адреса в NAT). Тому з вебсайту, якщо ведуться належні журнали, можна ідентифікувати користувача.

Застосування загальних інструментів криміналістичного аналізу й аналіз дампів. Місце встановлення основних компонентів браузера за замовчуванням зведено в табл. 3.

Таблиця 3.

Місце встановлення основних компонентів браузера Edge в ОС Windows

Компонент	Шлях
Основна програма	C:\Program Files (x86)\Microsoft\Edge\Application|Містить msedge.exe, бібліотеки, ресурси
Update Engine	C:\Program Files (x86)\Microsoft\EdgeUpdate|Служба оновлення браузера
User Data (профіль)	%LOCALAPPDATA%\Microsoft\Edge\User Data|Тут зберігаються всі профілі, історія, куки, логіни
Кеш	%LOCALAPPDATA%\Microsoft\Edge\User Data\Default\Cache|Chromium-кеш, IndexedDB, Media Cache
Розширення	%LOCALAPPDATA%\Microsoft\Edge\User Data\Default\Extensions\ Установлені розширення

Попри те, що виконавчий файл встановлюється в каталог для 32-бітних програм, це не обов'язково означає, що msedge буде саме 32-бітним. Звичайно в 64-бітній ОС встановлюється 64-бітний Edge та інші супутні компоненти, проте з використанням корпоративного інсталятора можна зробити різні вибори.

Формально незалежними компонентами від Edge, хоча вони встановлюються й оновлюються разом, є WebView2, widgets та xsocial. WebView2 – це портативна версія Edge, яка призначена для виконання в програмах як двигун графічного інтерфейсу. Інші дві програми – це новини з сайту MSN та для користувачів XBox. Обидві програми вимагають застосування WebView2.

64-бітний Edge застосовує наступні ключі реєстру – глобальні (HKLM), які визначають політики, версію, параметри оновлення, дозволи та користувача (HKCU), які зберігають налаштування користувача, синхронізацію, дозволи, стан профілю.

Програмні інтерфейси Windows для взаємодії з Edge приведені в табл. 4.

Таблиця 4.

Програмні інтерфейсі ОС Windows для взаємодії з Edge

API / Інтерфейс	Призначення
WinINET / WinHTTP	Edge використовує WinHTTP для системних запитів
ShellExecute / COM	Запуск браузера через msedge.exe з параметрами
Group Policy (ADMX)	Керування політиками через GPO
WebView2 Runtime	Вбудовування Edge як движка рендерингу в застосунки
DPAPI	Шифрування паролів у Login Data. Для розшифрування паролів потрібен SID користувача та ключі з Protect в реєстрі
Windows Search / Jump Lists	Інтеграція з пошуком, історією запуску

Реалізовані API оптимізовані під ОС Windows, що робить взаємодію та адміністрування під цією ОС зручнішим ніж, наприклад, з Google Chrome.

Інструменти загального призначення зведемо до табл. 5.

Таблиця 5.

Інструменти загального призначення, придатні для задач криміналістичного аналізу

№	Назва ПЗ	Призначення	Сайт
1	AccessData FTK Imager	Створення образів дисків й пам'яті для їх подальшого криміналістичного аналізу	exterro.com
2	Volatility	Інструмент для аналізу образів дисків і пам'яті, файлу підкачування, гібернації, інших дампів з великою кількістю плагінів для виконання всіх основних задач	github.com/ volatilityfoundation/ volatility
3	DB Browser for SQLite	Зручний графічний інтерфейс доступу до баз даних SqLite з можливістю виконувати довільні команди	sqlitebrowser.org
4	Janice	Програма для перегляду змісту JSON файлів як дерев, оптимізована для роботи з файлами великого розміру	github.com/Janice

продовження табл. 5.

5	ImHex	Багатофункціональна програма для перегляду бінарних файлів (для IndexedDB)	github.com/ WerWolv/ImHex
6	DCode	Програма для декодування різних форм запису дати/часу подій (підтримуються особливі формати зберігання Chromium), може застосовуватись до роботи з бінарними файлами	www.digital-detective.net/dcode/
7	CyberChef	Вебінструмент для декодування, дешифрування, парсингу Застосовується для декодування protobuf, base64, hex, gzip — вони часто зустрічаються в браузерних артефактах	gchq.github.io/ CyberChef
8	ShellBags Explorer	Дозволяє побічно підтвердити доступ до профілю браузера або його копіювання	tzworks.com
9	Registry Explorer	Перегляд реєстру Windows з криміналістичними функціями Може виявити налаштування браузера, політики, сліди запуску ericzimmerman.github.io	

Уявимо, що зроблено дамп активної операційної системи з браузером Edge за допомогою FTK Imager. Як отримати потрібні дані? Можна подивитись в базу даних в пам'яті. Операційна система для ефективності використовує механізм кешування файлів у пам'яті. Це означає, що коли Chrome читає файл History з диска, вміст цього файлу зберігається в оперативній пам'яті (у кеші файлової системи). Під час наступних звернень до файлу система може дуже швидко зчитати дані прямо з ОЗП, не звертаючись до повільного диска. Крім того, сам процес msedge.exe буде зберігати частини цієї БД у своїй власній пам'яті (кучі процесу), оскільки він працює з даними з файлу. В пам'яті буде знаходитись не обов'язково вся БД, але її значна й часто актуальна частина.

Коли створюється дамп пам'яті (наприклад за допомогою FTK Imager, Belkasoft Live RAM Capturer чи DumpIt), виконується наступне. Створюється знімок усієї оперативної пам'яті на даний момент часу, що включає. Дані всіх запущених процесів (виконавчий код, кучі, стеки тощо). Кеш файлової системи, в якому як раз можуть знаходитись блоки (сторінки) даних з файлу History й багатьох інших недавно використаних файлів. Різні інші структури даних ядра ОС.

Розглянемо більш детально застосування Volatility.

Після того, як створено дамп (наприклад в файл memory.raw) необхідно застосовувати спеціальні плагіни в профілі Win10x64.

Для версії 2.x починати треба з наступних плагінів

```
volatility -f memory.raw --profile=Win10x64 pslist
volatility -f memory.raw --profile=Win10x64 dlllist -p
volatility -f memory.raw --profile=Win10x64 cmdline -p
volatility -f memory.raw --profile=Win10x64 netscan
volatility -f memory.raw --profile=Win10x64 chromehistory
```

На що звертати увагу? Процеси msedge.exe — перевірити кількість, PID, час запуску. Відкриті вкладки — через chromehistory, cmdline, dlllist. Мережеві з’єднання — netscan покаже активні TCP/UDP сесії. Завантажені DLL — можуть вказати на розширення або шкідливі модулі. Кешовані URL / історія — витягується з ОЗП, навіть якщо профіль був очищений.

Як працює плагін chromehistory в версії 2.x? Сканує пам'ять - проходить по всьому дампу пам'яті, шукає сигнатури, структури, характерні для процесів Chrome. Знаходить SQLite БД – шукає в пам'яті фрагменти, які мають спеціальних заголовок. З окремих фрагментів в пам'яті намагається відновити цілісну структуру бази даних в пам'яті. Якщо це вдається, то до неї можна робити SQL-запити. Інші плагіни з префіксом “chrome” працюють аналогічно [10]. В версії 3.x концепція пошуку інша. ЇЇ використовує плагін Chrome HX [11]. Плагін НЕ сканує всю пам'ять навмання. Замість цього він використовує потужний механізм Volatility 3 під назвою yarascan. Yarascan

дозволяє шукати в пам'яті за певними шаблонами (YARA-правилами). У цьому разі плагін шукає не просто випадкові дані, а сигнатуру заголовка SQLite бази даних (перші байти файлу History). Коли Volatility знаходить такий заголовок, вона здатна визначити, чи є цей регіон пам'яті відображеним файлом (memory-mapped file). Операційна система, коли програма читає файл, може відобразити його вміст прямо у віртуальну пам'ять процесу. Це ефективний механізм доступу. Плагін витягує цей відображений файл цілком. По суті, він знаходить у пам'яті точну копію файлу History з диска (на момент створення дампа) і зберігає його у вигляді файлу на диск дослідника.

Цікавою є робота [12], в якій запропонована методологія для отримання даних з браузерів на платформі Chromium.

1. Ідентифікація класів і структур, які представляють браузер, вкладку, групу вкладок, відвідану URL-адресу тощо у вихідному коді проекту Chromium.

2. Генерування об'єктної структури таких класів і структур та сканування дампу пам'яті для пошуку об'єктів Browser як відправних точок аналізу пам'яті і, відповідно, виявлення дій користувача на основі Chromium.

3. Нарешті, виявивши об'єкт Browser, можна інтерпретувати значущі поля і дослідити всі доступні підоб'єкти, використовуючи змінні-покажчики, з метою вилучення значущої з точки зору криміналістичної експертизи інформації, наприклад, про відкриті вкладки і відвідані URL-адреси. Спрощено ідея роботи [22] полягає в тому, що програма мовою C++, скомпільована Visual C++ має певну структуру: не тільки типи даних стандартизовані в розмірах, але й структура класів, відношень між класами, формат переносу за слово за допомогою вказівників тощо. Тобто ми можемо відрізнити де йде клас, а де інший код. А значить ми можемо шукати структуру, яка пов'язана власне з класами, як вони реалізовані в конкретній редакції Chromium. Таким чином, необхідно обрати відповідну певному браузеру (Edge, Chrome тощо) версію Chromium й її скомпілювати в режимі збереження символів для відлагодження. Маючи символьні pdb файли ми можемо знайти в коді дампа місце, яке відповідає структурі полів класів. Структуру цих класів розробники браузерів на основі Chromium не змінюють.

Важливі класи та їх поля в коді Chromium приведені в табл. 6.

Таблиця 6.

Важливі класи та їх поля в коді Chromium

Клас	Поле	Коментар
Browser	profile_	Ім'я профілю чи профіль приватний
	tab_strip_model_	Вкладки й групи вкладок
	session_id_	ID об'єкта класу Browser
	bookmark_bar_state_	Чи показувати панель закладок
	window_has_shown_	Чи вікно показується
	user_title_	Заголовок вікна
ProfileImpl	path_	Шлях до AppData
TabStripModel	contents_data_	Перелік створених вкладок
group_model_	Дані про групи вкладок	
selection_model_	Індекс останньої використаної вкладки	
TabGroup	id_	ID групи вкладок
	visual_data_	Параметри відображення групи вкладок
	tab_count_	Кількість вкладок в групі вкладок
TabGroupId	token_	16-байтовий масив, який представляє унікальний ID
TabGroupVisualData	title_	Заголовок групи вкладок
	color_	Колір групи вкладок
NavigationControllerImpl	entries_	Перелік відвіданих URL
NavigationEntryImpl	frame_tree_	Дані, пов'язані з HTTP запитом
	unique_id_	ID запису
	title_	Заголовок відвіданої вебсторінки

продовження табл. 6.

	favicon_	Інформація про іконку сайта
	ssl_	дані SSL (сертифікат тощо)
	user_typed_url_	URL, що показується в адресному рядку
	timestamp_	Час відвідування (Chrome Epoch)
	http_status_code_	Код HTTP відповіді

Структурно ієрархію класів представимо на рис. 7.

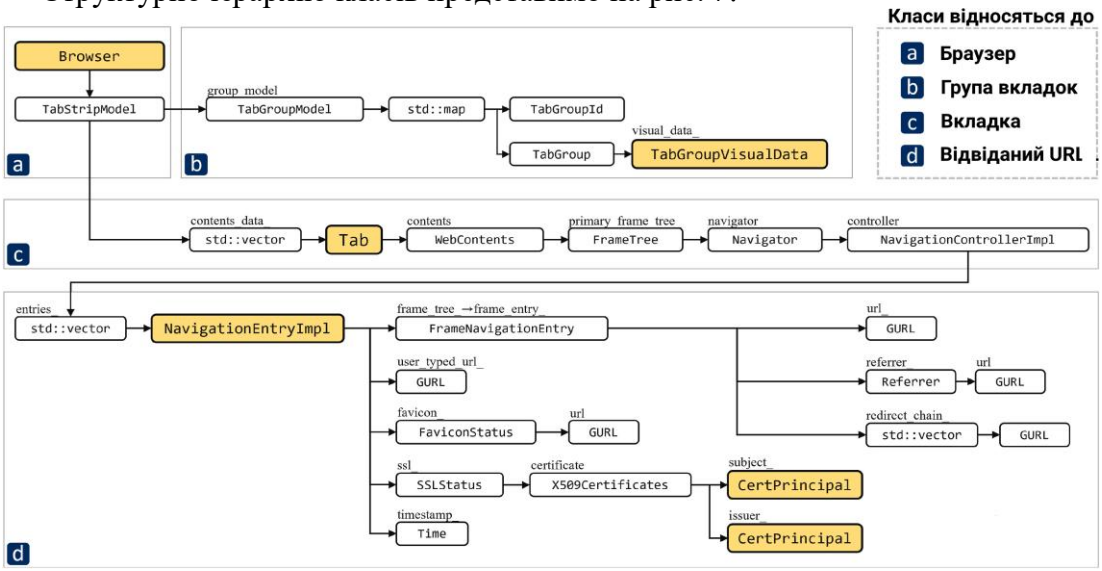


Рис. 7. Ієрархія ключових класів (адаптовано з [22])

Автори показали, що застосування алгоритму дозволяє визначити історію відвідувань за кожною вкладкою вікна. При чому не важливо, це приватний профіль чи звичайний. Для застосування підходу достатньо зробити дамп активного процесу msedge (через диспетчер задач чи програму procdump) без його зупинки й провести пошук за алгоритмом. Файл дампа вікна Edge звичайно не менше 0.5 Гб й більше 2 Гб якщо робити повний за допомогою procdump. Приклад знайденої частини, яка може бути заголовком вкладки в режимі Private показаний на рис. 8.

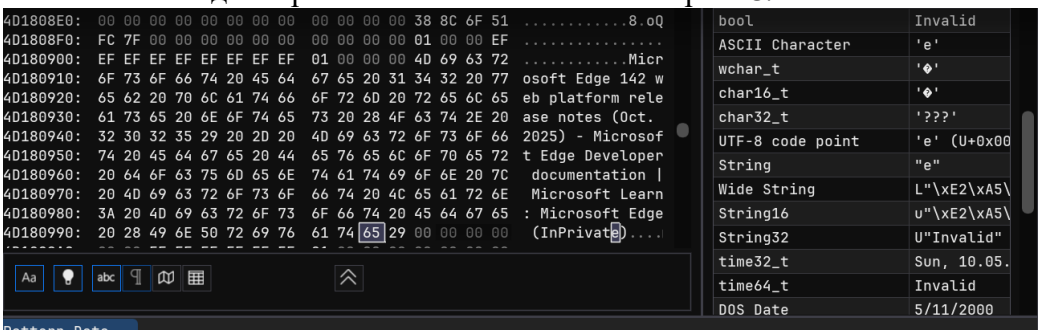


Рис. 8. Пошук в редакторі InPrivate текстового рядку в форматі UTF-8

Таким чином, загальні інструменти потенційно дозволяють більше ніж можна отримати з баз даних профіля. Проте це набагато складніше. Певний сенс вони мають лише коли наявний дамп, але не самі бази профіля або коли потрібно зібрати інформацію з дампу запущеного браузера з наявними вкладками в приватному режимі.

Застосування спеціалізованих засобів аналізу профілів браузерів на платформі Chromium. Далі розглянемо більш спеціалізовані інструменти.

BrowsingHistoryView від NirSoft [13]. Універсальна програма для перегляду історії . Достоїнством застосунку є те, що він може працювати майже з усіма

сучасними й не дуже браузерами під ОС Windows. Можливо зібрати URL, дату відвідування, кількість відвідувань й адресу, з якої потрапили на сторінку.

ChromeCacheView від NirSoft. Універсальний застосунок для перегляду кешу всіх браузерів на базі платформи Chromium. Можна провести вибірку й прочитати файл за його реальним шляхом в файловій системі.

ChromeCookieViewer від NirSoft. Універсальний застосунок для перегляду кукі всіх браузерів на базі платформи Chromium. Доступний зміст кукі, дата його створення, доступу та запланованого видалення.

Далі розглянемо інструменти аналізу, які реалізуються як плагіни до браузера. Тобто збирають дані з середини. Їх перевага – кросплатформовість, відкритий код (незалежно від ліцензії).

WebHistorian: Education edition [14]. Плагін дозволяє переглядати сайти, пошукові запити та час відвідувань. Історія відвідувань демонструється за допомогою спеціального графіку, показаного на рис. 9. У кожне коло можна перейти за детальною інформацією й переглянути таблицю по відвідуванням сайту: сторінка, заголовок й час.

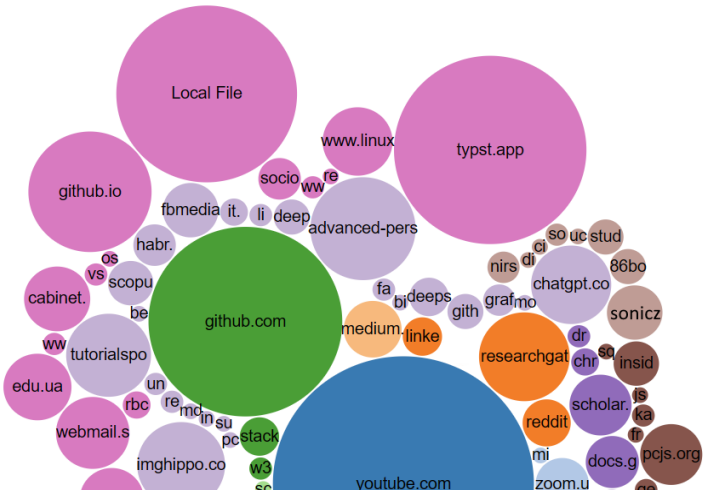


Рис. 9. Приклад візуалізації відвідуваності та сайтів за відвідуваністю (розмір) та категорією (колір)

Історія пошукових запитів візуалізується як хмара тегів за окремими словами. Але по кожному слову можна клацнути правою кнопкою миші й переглянути запити, які містили це ключове слово. Кольорова карта візуалізує активність користувача за годинами та днями тижня. Копія екрану показана на рис. 10. Знизу кольорова легенда показує яка приблизно кількість візитів спостерігається за годину. По кожному квадратику можна клацнути, та переглянути які конкретно сторінки були переглянуті в цей період часу.

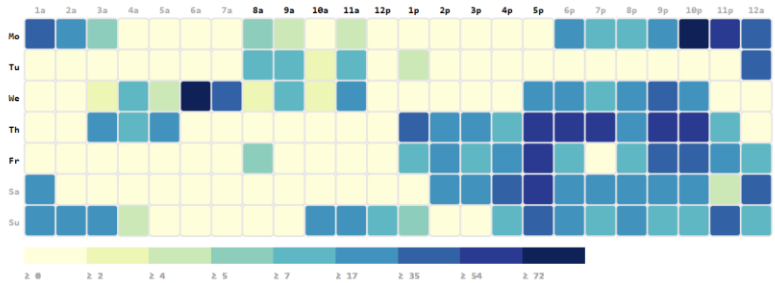


Рис. 10. Приклад візуалізації відвідуваності за годинами та днями тижня

History Trends Unlimited. Аналогічний до попереднього плагін [15], але менш зручний в користуванні, оскільки представляє дані просто як таблиці. Проте, в нього є можливість класифікації за типом переходу (пошуковий запит, пряме введення адреси, перехід з іншої адреси, згенероване ШІ посилання тощо) – рис. 11.

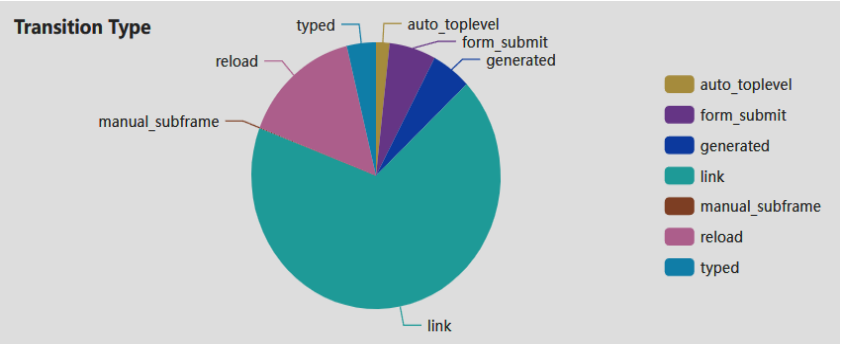


Рис. 11. Кругова діаграма за типами переходів [16].

Нарешті, розглянемо спеціалізовану програму для комплексного збору даних, яка називається Hindsight [17]. Ця програма реалізується в двох версія й по суті являє собою Python скрипт. В графічній версії інтерфейс реалізується як вебінтерфейс. В ньому можна задати шлях до профілю, запустити перегляд, переглянути загальну статистику та експортувати результати аналізу в Excel файл чи в JSON. Приклад загальних результатів аналізу показаний на рис. 12.

Detected Chrome version:	130-134	Login Data records:	202	DIPS Items:	0
URL records:	2570	Preference Items:	9582	Extension Rules records:	297
Download records:	63	Extensions:	21	Extension Scripts records:	32022
Cache records:	4920	Extension Cookie records:	2	Extension State records:	10512
GPU Cache records:	0	IndexedDB records:	9760	Local Extension Settings records:	57076
Cookie records:	6136	Session Storage records:	163928	Managed Extension Settings records:	0
Local Storage records:	28703	Site Characteristics records:	42812	Sync App Settings records:	0
Bookmark records:	52	File System Items:	495	Sync Extension Settings records:	465
Autofill records:	830	DIPS Popup Items:	32	HSTS records:	3961

Рис. 12. Статистика обробки профілю Hindsight

В порівнянні з попередніми інструментами цей інструмент найменш зручний, оскільки не має фактично графічного інтерфейсу, який допомагає аналізувати дані. Проте, Hindsight надає певну інформацію, яку не можливо отримати за допомогою інших інструментів. Це інформація про те, на яких сайтах застосовувалось автозаповнення форм (зокрема з паролями) та під яким логіном проводилась авторизація, це всі збережені дані у storage для вебзастосунків й це всі запити, які проводять періодично розширення. У цілому програма Hindsight – це гарна точка відліку для розробки застосунків з певними алгоритмами аналізу та візуалізації.

Висновки. Розглянуто аналіз збережених в браузері даних, які мають бути потрібні як докази при криміналістичному аналізі. Детально проаналізована загальна структура та формат збереження даних в профілі браузера. Описане застосування загальних інструментів криміналістичного аналізу й аналіз дампів. Проаналізоване застосування спеціалізованих засобів аналізу профілів браузерів на платформі Chromium, а саме такі.

1. Спеціалізовані утиліти для збору даних з позиції одного типу інформації (кеш, історія відвідувань, кукі тощо). Певна інформація вимагає закриття браузера, інша може бути зібрана й при відкритому браузері.
2. Комплексні засоби аналізу, реалізовані як плагіни для браузера. При цьому підході до реалізації системи не до всієї інформації наявний доступ, оскільки наявні обмеження браузера для плагінів, введені з метою безпеки й не тільки.
3. Комплексні засоби аналізу, яким не потрібен запуск в межах браузера. Ці інструменти здатні зібрати найбільшу кількість інформації. Проте звичайно працювати з нею не дуже зручно.

Список літератури

1. Mandine L. Browser Forensics. Medium. 2024. URL: https://medium.com/@laurent.mandine/browser-forensics-89429fe0749f

2. Кіберпсихологія у вимірах сучасного наукового дискурсу: монографія / С. Хаджирадєва, Ю. Левін, М. Тодорова; ред. С. К. Хаджирадєва. Одеса: Астропринт, 2024. 240 с.
3. The official GitHub of the Chromium. URL:https://github.com/chromium/chromium
4. Dušek D. Identification of specific Google Chrome user based on analysis of its application data. *Proc. of Studentská Konference Inovací, Technologií a Vědy v IT*. *Brno*. 2016. URL: https://excel.fit.vutbr.cz/submissions/2016/044/44.pdf
5. Shafqat N. Forensic investigation of user's web activity on Google Chrome using various forensic tools. *Int. J. Comput. Sci. Netw. Secur*. 2016. No.16(9). P.123–132.
6. Rathod D. Web browser forensics: Google Chrome. *Int. J. of Advanced Research in Computer Science*. 2017. V.8. No. 7. URL: https://ijarcs.info/index.php/Ijarcs/article/view/4433
7. Flowers C., Haider A. Web browser artefacts in private and portable modes: A forensic investigation. *International Journal of Electronic Security and Digital Forensics*. 2016. No.8. P.99–117. URL: https://doi.org/10.1504/IJESDF.2016.075583
8. Horsman G., Findlay B., Edwick J., Asquith A., Swannell K. A forensic examination of web browser privacy-modes. Forensic Science International: Reports. Volume 1. 2019. URL: https://doi.org/10.1016/j.fsir.2019.100036.
9. Hughes K., Papadopoulos P., Pitropakis N., Smales A. Private Mode: Is It What We Were Promised? *Computers.* 2021. No. 10. P.165. URL: https://doi.org/10.3390/computers10120165
10. Volatility plugin: Chrome History. URL: https://blog.superponible.com/2014/08/31/volatility-plugin-chrome-history/
11. Chrome history plugin for Volatility 3. URL: https://github.com/its-radio/volatility_plugins/blob /main/chrome_hx/chrome_hx.py
12. Choi G., Bang J., Lee S., Park J. Chracer: Memory analysis of Chromium-based browsers. *Forensic Science International: Digital Investigation.* 2023. Vol. 46. DOI: 10.1016/j.fsidi.2023.301613.
13. Unique collection of small and useful freeware utilities. URL: nirsoft. net
14. Digital Forensic Analysis of Web Browser Records. URL: webhistorian.com
15. History Trends Unlimited https://chromewebstore.google.com/detail/history-trends-unlimited/pnmchffiealhkdloeffcdnbgdnedheme
16. Chromium is an open-source browser project. URL: https://chromium.googlesource.com/chromium/src/+/master/chrome/common/extensions/api/history.json
17. Hindsight: Internet history forensics for Google Chrome/Chromium. URL: https://github.com/obsidianforensics/hindsight

WEB BROWSER HISTORY FORENSIC ANALYSIS

M.V. Vidin[1], O.A. Stopakevych[1], A.O. Stopakevych[2]

[1]National Odesa Polytechnic University
1, Shevchenko Ave., Odesa, 65044, Ukraine
[2]State University of Intellectual Technologies and Telecommunications
1, Kuznechna Str., Odesa, 65023, Ukraine
Email: stopakevich@gmail.com

The present study analyzes the capabilities of existing software tools for the forensic analysis of the history of the most commonly used browsers based on the Chromium platform (Chrome, Edge, Opera, etc.) to determine their effectiveness. The existing scheme for storing user data and actions in the Edge profile has been formalized and analyzed, and the locations where various settings and files of this browser are stored have been identified. A thorough analysis of the Chromium source code and the actual data stored by the Edge browser was conducted to elucidate the relationships and interconnections between the entities of the SQLite database of the profile and the contents of its main JSON file. The existing software that facilitates forensic analysis of profile data is reviewed. The following are utilities for working with data and dumps, specialized utilities for collecting data from a browser profile from the perspective of one type of information—cache, visit history, cookies—and more complex tools for analyzing browser profiles. The latter category facilitates the collection of a more substantial amount of information. However, in general, the potential exists for significant improvement in the realm of complex software, given its present suboptimal functionality, particularly from the perspective of forensic experts who utilize it.

Keywords: forensic analysis, browser, data collection, browser plugins, activity analysis, malware, criminal activity timeline.

DOI 10.15276/imms.v15.no4.536
УДК 004.056.53:004.912

Informatics and Mathematical Methods in Simulation
Vol.15 (2025), No. 4, pp. 536-544

СИСТЕМА АВТОМАТИЗОВАНОГО АНАЛІЗУ ЦИФРОВОГО СЛІДУ КОРИСТУВАЧА В СОЦІАЛЬНИХ МЕДІА З ВИКОРИСТАННЯМ OSINT

А. В. Власова, В. О. Назаров, І. А. Ярова, Н. І. Кушніренко

Національний університет «Одеська Політехніка»
1, Шевченка пр., Одеса, 65044, Україна
Email: kushnirenko@op.edu.ua

У статті представлено систему автоматизованого аналізу цифрового сліду користувача в соціальних медіа з використанням підходів OSINT, адаптованого для української мови. Зростання активності в соціальних мережах створює ризики для приватності через накопичення персональної інформації. Метою роботи є створення спеціалізованої системи для комплексного аналізу цифрової присутності користувачів у Telegram та YouTube із урахуванням морфологічних особливостей української мови та культурного контексту. Проведено аналіз існуючих методів дослідження соціальних платформ, виявлено обмеження при роботі з українським контентом через недостатню точність розпізнавання мови та відсутність спеціалізованих лінгвістичних моделей. Запропонована система складається з трьох послідовних алгоритмів: алгоритму збору даних через API Telegram та YouTube, алгоритму комплексного аналізу текстового контенту з використанням TF-IDF для виділення ключових термінів та langdetect для автоматичного визначення мови, алгоритму побудови інтегрованого профілю з нормалізацією різнотипних даних на основі текстового, соціального, часового та поведінкового компонентів. Експериментальне тестування на вибірці зі ста користувачів різних вікових категорій показало точність системи 73% для українського контенту, що на 18-25% перевищує міжнародні аналоги при середньому часі обробки 85-95 секунд на користувача. Значення F1-score становлять 0.79 для збору даних, 0.66 для алгоритму текстового аналізу та 0.64 для алгоритму побудови профілю. Наукова новизна полягає у створенні спеціалізованої системи для української мови з урахуванням її морфологічних особливостей. Результати можуть використовуватися спеціалістами з кібербезпеки для аудиту цифрового сліду, дослідниками соціальних мереж для аналізу поведінкових патернів та фахівцями OSINT для верифікації інформації з відкритих джерел.
Ключові слова: OSINT, цифровий слід, соціальні медіа, Telegram, YouTube, машинне навчання, аналіз тексту.

Вступ. Зростання активності користувачів у соціальних мережах призводить до формування значних обсягів цифрових слідів, які містять персональну інформацію про їх власників. Open Source Intelligence являє собою дисципліну збору та аналізу інформації з публічно доступних джерел для отримання розвідувальних висновків. Хоча методи роботи з відкритими джерелами використовувались протягом століть, сучасне розуміння OSINT сформувалось з розвитком інтернету та цифрових комунікацій.

Цифровий слід користувача формується через взаємодію з різноманітними онлайн-платформами та сервісами. Цей слід включає як свідомо залишені дані, такі як публікації в соціальних мережах, так і несвідомо створену інформацію, включаючи метадані, часові відмітки активності та патерни поведінки [1]. Дослідження показують, що навіть обмежений набір даних з соціальних мереж може розкрити значну кількість персональної інформації, що створює як можливості для легітимного аналізу, так і серйозні питання щодо приватності [2].

Методи аналізу соціальних мереж характеризуються різноманітністю підходів. Мережевий аналіз зосереджується на структурі зв'язків між користувачами та поширенні інформації. Контентний аналіз досліджує текстову та мультимедійну інформацію через сентимент-аналіз, тематичне моделювання та класифікацію тексту [3]. Поведінковий

аналіз вивчає патерни активності, часові характеристики та геопросторовий контекст. Інтегровані підходи поєднують мультимодальний аналіз та застосування машинного навчання для комплексного дослідження користувацької активності [4].

Сучасні дослідження все частіше використовують техніки машинного навчання для автоматизації процесу аналізу [5]. Алгоритми класифікації дозволяють категоризувати контент за тематикою чи тональністю. Нейронні мережі можуть обробляти складні мультимодальні дані, включаючи текст, зображення та відео одночасно. Особливого розвитку набули методи обробки природної мови для аналізу текстового контенту [6].

Telegram займає унікальну нішу серед месенджерів через поєднання приватного спілкування з публічними каналами та групами. Архітектура включає приватні чати, групи, супергрупи та канали. Особливістю Telegram є система пересилання повідомлень з збереженням інформації про оригінальне джерело, що створює можливості для відстеження поширення інформації між каналами [7]. YouTube представляє іншу модель цифрової присутності, орієнтовану на довготривалий відеоконтент. Платформа надає багатий набір метаданих для кожного відео, включаючи статистику переглядів, коментарі та інформацію про плейлисти [8].

Основними обмеженнями існуючих OSINT інструментів є орієнтація на англомовний контент, складність налаштування та використання, а також фрагментарність функціональності. Більшість рішень вирішують окремі завдання без можливості комплексного аналізу цифрового сліду користувача. Для української мови це особливо актуально через складну морфологію та відсутність великих анотованих корпусів даних [9].

Мета роботи. Метою даної роботи є розробка системи автоматизованого аналізу цифрового сліду користувача в соціальних медіа з використанням підходів OSINT, адаптованої для української мови та культурного контексту.

Для досягнення поставленої мети необхідно розв’язати такі завдання:

- розробити алгоритм збору публічно доступних даних з соціальних медіа;
- розробити алгоритм комплексного аналізу текстового контенту з урахуванням специфіки української мови;
- розробити алгоритм побудови інтегрованого цифрового профілю користувача;
- реалізувати зазначені алгоритми у складі єдиної системи;
- провести тестування розробленої системи та проаналізувати її ефективність.

Основна частина. Вибір Telegram та YouTube обумовлений їх популярністю серед української аудиторії та наявністю стабільних API для збору публічно доступних даних з каналів, груп, відео та коментарів.

Алгоритм збору даних реалізовано у вигляді послідовного процесу з п'яти основних етапів. Перший етап включає ініціалізацію та ідентифікацію користувача. Алгоритм отримує початкові дані користувача, такі як username або ID профілю, та перевіряє доступність профілю на обох платформах. Для Telegram використовується Telegram Bot API через офіційний клієнт, а для YouTube застосовується Data API v3. Другий етап передбачає збір базової інформації профілю користувача, включаючи ім'я користувача, опис профілю, аватар, дату створення акаунту, кількість підписників та підписок. Ця інформація формує базову структуру цифрового профілю користувача та служить основою для подальшого аналізу. Третій етап зосереджується на аналізі контентної активності користувача. Алгоритм збирає всі доступні публічні публікації, коментарі та реакції користувача. Для Telegram аналізуються повідомлення у публічних каналах та групах, включаючи переслані повідомлення, до яких користувач має доступ або є учасником. Для YouTube збираються відео користувача, його коментарі під власними та чужими відео, створені плейлисти та їх наповнення. Четвертий етап включає побудову соціального графу користувача через аналіз підписок на канали та користувачів, взаємодій з іншими учасниками спільнот, частоти спілкування з

конкретними контактами. П'ятий етап передбачає збір метаданих та темпоральної інформації, включаючи часові мітки активності, інформацію про пристрої користувача та інші технічні метадані.

На основі розробленого алгоритму створено програмний застосунок, який реалізує всі описані етапи збору та обробки даних. Для забезпечення стабільної роботи застосунку передбачено систему обробки помилок. При тимчасовій недоступності API платформи алгоритм переходить в режим очікування з поступовим збільшенням інтервалу між спробами підключення згідно з формулою:

$$t_{wait} = t_0 * 2^n \quad (1)$$

де t_0= 1 секунда, n – номер спроби підключення. Також реалізовано локальний кеш для оптимізації продуктивності та зменшення навантаження на зовнішні API. Час життя кешу визначається типом даних: 300 секунд для профільних даних користувачів, 60 секунд для динамічного контенту (публікації, коментарі) та 900 секунд для статичних метаданих (інформація про канали, групи). Така диференціація забезпечує баланс між актуальністю інформації та ефективним використанням обмежених квот API платформ.

Реалізація етапів 2-5 має специфічні особливості для кожної платформи. Адаптація алгоритму для Telegram включає роботу з месенджер-специфічними функціями через Telegram Bot API. Алгоритм використовує метод getChat для отримання інформації про публічні канали та групи, включаючи назву, опис, кількість учасників та іншу базову інформацію. Метод getChatMember дозволяє аналізувати статус користувача в публічних спільнотах, визначати його роль та рівень активності. Для отримання візуальної інформації застосовується метод getUserProfilePhotos, який забезпечує доступ до аватарів користувачів. Окрему увагу приділено обробці пересланих повідомлень, оскільки Telegram зберігає метадані про оригінальне джерело, що дозволяє відстежувати ланцюжки поширення інформації між каналами. Алгоритм також враховує специфіку публічних та приватних груп, адаптуючи методи збору даних відповідно до рівня доступності інформації.

Для YouTube адаптація алгоритму зосереджується на комплексному аналізі публічних каналів користувачів через YouTube Data API v3. Процес збору даних починається з методу channels.list, який надає базову статистику каналу: кількість підписників, загальну кількість переглядів, дату створення та опис каналу. Метод search.list використовується для отримання списку всіх публічних відео користувача з можливістю фільтрації за датою публікації та типом контенту. Для кожного відео алгоритм збирає детальні метадані через метод videos.list, включаючи назву, опис, теги, тривалість, статистику переглядів, лайків та дизлайків. Особливу цінність представляє аналіз коментарів через метод commentThreads.list, який дозволяє вивчати не тільки коментарі користувача під власними відео, але й його активність під відео інших авторів. Алгоритм також аналізує створені користувачем плейлисти через метод playlists.list та їх наповнення через playlistItems.list, що розкриває тематичні інтереси та патерни споживання контенту. Додатково враховуються дані про підписки користувача на інші канали, що формує уявлення про його соціальне оточення та сфери інтересів у відеоконтенті.

Розроблений алгоритм збору даних реалізує багаторівневу систему кешування для оптимізації продуктивності. Система обробки помилок адаптована до особливостей кожної платформи: Telegram Bot API має обмеження на кількість запитів та специфічні коди помилок, які потребують окремої обробки, а YouTube API має систему квот з лімітом 10000 одиниць на день, що вимагає ретельного планування запитів та оптимізації використання API ресурсів.

Зібрані дані з обох платформ обробляються незалежно та передаються до наступного етапу аналізу, де вони об'єднуються на рівні побудови інтегрованого профілю користувача. Для Telegram аналізуються доступні повідомлення користувача у публічних каналах та групах, частота активності, типи контенту який він публікує або

пересилає. Особлива увага приділяється аналізу форвардинга повідомлень, що може розкрити мережу інформаційних джерел користувача та його уподобання. Для YouTube адаптація зосереджується на аналізі публічних каналів користувачів, їх відео контенту, коментарів, плейлистів та підписок на інші канали. YouTube API дозволяє отримувати детальну статистику каналу, включаючи кількість переглядів, підписників, лайків та коментарів, що надає можливість оцінити рівень впливу користувача на платформі. Алгоритм аналізує метадані відео, включаючи заголовки, описи, теги, тривалість та дату публікації для побудови профілю інтересів користувача.

Текстовий аналіз займає центральне місце в системах OSINT, оскільки мовна поведінка користувачів розкриває найбільше інформації про їх особистісні характеристики, світогляд та соціальні зв'язки. Розроблений алгоритм поєднує традиційні методи обробки природної мови з сучасними підходами частотного аналізу, адаптованими для української мови.

Алгоритм текстового аналізу складається з п'яти етапів обробки. Перший етап включає нормалізацію та очищення тексту від технічних артефактів, HTML тегів, спеціальних символів та інших елементів, які не несуть семантичного навантаження. Процес нормалізації включає приведення тексту до стандартного формату UTF-8, корекцію кодування символів, обробку емоджі та спеціальних символів. Другий етап передбачає токенізацію та лематизацію тексту з урахуванням морфологічних особливостей української мови. Використовується спеціалізований токенізатор, який враховує українські конструкції, прийменники з апострофом, складні числівники та назви власні. Третій етап включає автоматичне визначення мови тексту з використанням бібліотеки langdetect, що важливо для багатомовного контенту користувачів. Алгоритм використовує статистичний підхід на основі частотного аналізу символьних n-грам для української та англійської мов. Визначення мови дозволяє коректно застосовувати відповідні лінгвістичні ресурси: для української мови використовуються спеціалізовані словники та токенізатори, адаптовані до морфологічних особливостей української мови, тоді як для англійської застосовуються стандартні бібліотеки обробки природної мови. Четвертий етап алгоритму реалізує сентимент-аналіз текстового контенту для визначення емоційного забарвлення повідомлень користувача. Використовується гібридний підхід, який поєднує словникові методи з елементами частотного аналізу. Алгоритм класифікує тексти за трьома основними категоріями емоційного забарвлення: позитивне, негативне та нейтральне. П'ятий етап включає частотний аналіз ключових слів та виділення сутностей з тексту користувача. Алгоритм використовує метод TF-IDF для автоматичного виявлення найбільш важливих термінів у корпусі текстів користувача. Вага терміну *t* у конкретному тексті *d* розраховується за формулою:

$$TF - IDF(t,d) = (f(t,d) / max\_freq(d)) * log(N / |\{d \in D : t \in d\}|) \quad (2)$$

де *f(t,d)* - частота терміну t у тексті d, *max_freq(d*) - максимальна частота будь-якого терміну в тексті d, *N* - загальна кількість текстових фрагментів користувача, $|\{d \in D : t \in d\}|$ - кількість текстових фрагментів, що містять термін t.

Також реалізовано обробку неструктурованого тексту з урахуванням інтернет-сленгу, скорочень та неологізмів, характерних для української інтернет-культури. Створено спеціалізований словник з понад 800 популярних українських скорочень з їх розшифровками та контекстуальними значеннями. Це суттєво підвищує точність аналізу сучасного українського цифрового контенту та зменшує кількість помилок при обробці нестандартних текстів.

Побудова комплексного цифрового профілю користувача є третім алгоритмом системи аналізу, який інтегрує результати всіх попередніх стадій обробки у єдину структуровану модель. Алгоритм побудови профілю функціонує через два основних етапи. Початковий етап включає нормалізацію та стандартизацію всіх зібраних даних до єдиної системи координат. Процес нормалізації передбачає приведення всіх метрик до стандартизованої шкали [0,1] за допомогою функції мін-макс нормалізації:

$$x\_norm = (x - x\_min) / (x\_max - x\_min) \quad (3)$$

де x - початкове значення, x_min та x_max - мінімальне та максимальне значення в діапазоні.

Соціальний компонент характеризує мережу соціальних зв'язків користувача, рівень його впливу через співвідношення підписників до підписок, та активність взаємодії з іншими користувачами через кількість коментарів, відповідей та реакцій на публікації інших користувачів. Часовий компонент відображає патерни активності користувача в часі, включаючи періоди найвищої активності, регулярність публікацій та зміни в поведінці протягом різних періодів. Поведінковий компонент аналізує типи контенту, які користувач споживає та створює через категоризацію за форматом (текст, відео, зображення) та тематикою, його реакції та взаємодії з різними типами матеріалів через частоту лайків, коментарів, репостів та переглядів. Другий етап передбачає розрахунок базових компонентів цифрового профілю. Текстовий компонент включає аналіз лексичного багатства користувача, визначення домінуючих тематик на основі частотного аналізу ключових термінів, оцінку емоційного профілю через співвідношення позитивних, негативних та нейтральних повідомлень, та аналіз стилістичних особливостей через довжину речень і використання специфічної лексики. Лексичне багатство розраховується за формулою TTR:

$$TTR = |V| / N \quad (4)$$

де $|V|$ - кількість унікальних слів, N - загальна кількість слів у корпусі текстів користувача.

Результати. Оцінка ефективності розробленої системи проводилась на публічно доступних профілях користувачів Telegram та YouTube різних вікових категорій та рівнів активності. Тестування проводилось на 100 користувачах, профілі яких були відібрані з публічних каналів та груп української тематики. Структура тестової вибірки представлена в табл. 1, яка демонструє розподіл учасників дослідження за віковими групами та рівнем активності на досліджуваних платформах.

Таблиця 1

Характеристики тестової вибірки користувачів

Віковий діапазон	Кількість користувачів	Рівень активності	Telegram	YouTube
18-25 років	36 (36%)	Високий	30	24
26-35 років	40 (40%)	Помірний	36	32
36-50 років	24 (24%)	Низький	20	16
Загалом	100 (100%)	-	86	72

Алгоритм побудови цифрового профілю показує варіативну точність залежно від повноти вхідних даних. Серед 100 досліджуваних користувачів 36 мали високу активність переважно у віці 18-25 років, демонструючи точність профілювання 78%. Користувачі з помірною активністю, які становили 40% вибірки переважно у віці 26-35 років, показали точність 66%. Користувачі з низькою активністю, переважно старші за 36 років і складаючи 24% вибірки, демонстрували точність профілювання 59%.

Аналіз продуктивності алгоритму збору даних показав середній час обробки одного користувача приблизно 85 секунд для Telegram та 95 секунд для YouTube. Дослідження точності алгоритмів текстового аналізу показало високі результати для української мови: визначення мови досягло точності 91%, сентимент-аналіз 74%, тематична класифікація 69%, виділення ключових слів 77%. Для англійської мови результати становили відповідно 92%, 71%, 67%, 74%. Змішаний контент демонстрував дещо нижчі показники: 87%, 65%, 64%, 69%. Особливі виклики для системи становлять тексти з високим рівнем іронії, сарказму або використанням сленгу, де точність аналізу

знижується, що пов'язано зі складністю автоматичного розпізнавання контекстуальних та культурних нюансів української інтернет-комунікації.

Порівняльний аналіз з існуючими OSINT інструментами представлено в табл. 2.

Таблиця 2

Аналіз ефективності системи відносно існуючих рішень

Система	Точність для української мови	Швидкість обробки	Платформи	Спеціалізація
Maltego	55%	180-240 сек	Багатоплатформна	Міжнародна
OSINT Framework	48%	200-280 сек	Багатоплатформна	Міжнародна
Social-Analyzer	52%	190-260 сек	Соціальні мережі	Міжнародна
Розроблена система	73%	85-95 сек	Telegram, YouTube	Українська

Результати показують переваги розробленої системи у спеціалізації на українському контенті, що забезпечує приріст точності на 18-25% порівняно з універсальними міжнародними рішеннями. Показники точності для систем-аналогів отримані шляхом тестування на власній вибірці з 50 українських профілів. Різниця у швидкості обробки пояснюється різним функціоналом систем: універсальні інструменти виконують ширше коло завдань, тоді як розроблена система оптимізована для конкретних платформ Telegram та YouTube.

Функція оцінки якості результатів базується на метриці F1-score:

$$F1 = 2*(precision*recall)/(precision+recall) \quad (5)$$

де $precision = TP/(TP+FP)$, $recall = TP/(TP+FN)$, TP - true positive, FP - false positive, FN - false negative. Середнє значення F1-score для трьох алгоритмів системи становить 0.79 для збору даних, 0.66 для текстового аналізу, 0.64 для побудови профілю, що свідчить про задовільну якість роботи алгоритмів з урахуванням складності завдань автоматичного аналізу соціальних медіа.

Аналіз помилок всієї системи в цілому виявив основні джерела неточностей, які представлені на рис. 1.

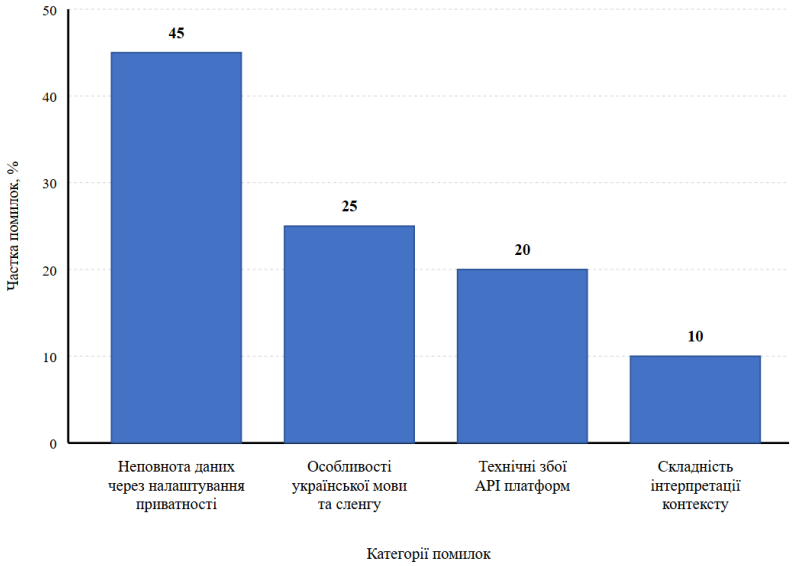


Рис. 1. Гістограма помилок системи

Найбільшу частку становлять проблеми з неповнотою даних через налаштування приватності користувачів (45%), особливості української мови та інтернет-сленгу (25%), технічні збої API платформ (20%) та складність інтерпретації контекстуальних значень (10%). Розподіл причин помилок показує, що найбільший вплив на якість аналізу має доступність даних, тоді як технічні та лінгвістичні виклики становлять менший, але значущий внесок у загальну похибку системи. Оцінка точності побудови комплексного

цифрового профілю проводилась шляхом порівняння результатів роботи системи з експертними оцінками (табл. 3).

Таблиця 3

Точність ідентифікації характеристик цифрового профілю

Характеристика	Точність	Кількість зразків
Основні інтереси користувача	75%	100
Емоційний профіль	70%	100
Рівень соціальної активності	77%	100
Професійна сфера діяльності	65%	87
Часові патерни активності	87%	76

Результати показують найвищу точність у визначенні часових патернів активності (87%), що пояснюється об'єктивним характером темпоральних даних та їх меншою залежністю від інтерпретації. Рівень соціальної активності визначається з точністю 77%, оскільки метрики підписок, коментарів та взаємодій піддаються прямому кількісному аналізу. Основні інтереси користувача ідентифікуються з точністю 75% завдяки застосуванню методу TF-IDF для виділення ключових термінів у текстовому контенті. Емоційний профіль визначається з точністю 70%, що є задовільним результатом з урахуванням складності сентимент-аналізу для української мови. Найнижчу точність демонструє визначення професійної сфери діяльності (65%), оскільки ця характеристика часто не виражена в публічному контенті користувачів.

Загальна точність системи для української мови розраховувалась як середньозважене значення точності визначення п'яти ключових характеристик користувача з таблиці 3: основні інтереси, емоційний профіль, рівень соціальної активності, професійна сфера та часові патерни. Середнє арифметичне цих показників становить 73% для користувачів з помірною активністю, що і використовується як базова метрика порівняння системи з аналогами.

Тестування стабільності роботи алгоритмів при різних умовах експлуатації показало задовільну стійкість до тимчасової недоступності API соціальних платформ. Система автоматичного відновлення з використанням експоненційного backoff забезпечує продовження роботи після усунення технічних проблем з відновленням 89% функціональності протягом 3-5 хвилин. При збільшенні навантаження система демонструє лінійне зростання часу обробки, зберігаючи можливість паралельної обробки до 5 профілів одночасно без суттєвого зниження продуктивності.

Дослідження масштабованості розробленої системи показало, що система ефективно справляється з обробкою користувачів різного рівня активності. Для користувачів з високою активністю (понад 500 публікацій) середній час обробки збільшується до 120-140 секунд, проте точність результатів зростає до 82% завдяки більшому обсягу даних для аналізу. Користувачі з помірною активністю (100-500 публікацій) обробляються за стандартний час 85-95 секунд з точністю 73%. Для користувачів з низькою активністю (менше 100 публікацій) час обробки скорочується до 60-70 секунд, однак точність знижується до 64% через обмежений обсяг вхідних даних.

Висновки. Розроблено систему автоматизованого аналізу цифрового сліду користувача в соціальних медіа з використанням підходів відкритого збору інформації, адаптовану для української мови та культурного контексту. Проведено комплексний аналіз існуючих методів дослідження соціальних платформ та виявлено їх основні обмеження при роботі з україномовним контентом. Розроблено алгоритм збору публічно доступних даних, який включає п'ять основних етапів та забезпечує систематичний підхід до отримання релевантних даних при дотриманні технічних обмежень. Реалізовано спеціалізовані адаптації для роботи з Telegram Bot API та YouTube Data API v3. Запропоновано алгоритм комплексного аналізу текстового контенту з використанням методу TF-IDF та бібліотеки langdetect для автоматичного визначення мови, адаптовану для української мови. Розроблено алгоритм побудови інтегрованого цифрового профілю,

який включає текстовий, соціальний, часовий та поведінковий компоненти. Експериментальне тестування підтвердило ефективність системи з точністю 73% для українського контенту, що на 18-25% перевищує показники міжнародних аналогів при середньому часі обробки 85-95 секунд на користувача.

Список літератури

1. Azucar D., Marengo D., Settanni M. Predicting the Big 5 personality traits from digital footprints on social media: A meta-analysis. *Personality and Individual Differences.* 2018. Vol. 124. P. 150-159.
2. Feher K. Digital identity and the online self: Footprint strategies – An exploratory and comparative research study. *Journal of Information Science.* 2019. P. 016555151987970.
3. Dang N. C., Moreno-García M. N., De la Prieta F. Sentiment Analysis Based on Deep Learning: A Comparative Study. *Electronics.* 2020. Vol. 9. No. 3. P. 483.
4. Explainable AI for Psychological Profiling from Behavioral Data: An Application to Big Five Personality Predictions from Financial Transaction Records / Y. Ramon et al. *Information.* 2021. Vol. 12. No. 12. P. 518.
5. Nandwani P., Verma R. A review on sentiment analysis and emotion detection from text. Social Network Analysis and Mining. 2021. Vol. 11, no. 1.
6. Medhat W., Hassan A., Korashy H. Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*. 2014. Vol. 5. No. 4. P. 1093-1113.
7. Predicting Consumers' Decision-Making Styles by Analyzing Digital Footprints on Facebook / Y.-J. Chen et al. *International Journal of Information Technology & Decision Making.* 2019. Vol. 18. No. 02. P. 601-627.
8. Predicting Loneliness through Digital Footprints on Google and YouTube / E. Ahmed et al. *Electronics.* 2023. Vol. 12. No. 23. P. 4821.
9. Deeva I. Computational Personality Prediction Based on Digital Footprint of A Social Media User. *Procedia Computer Science*. 2019. Vol. 156. P. 185-193.

SYSTEM FOR AUTOMATED ANALYSIS OF USER DIGITAL FOOTPRINT IN SOCIAL MEDIA USING OSINT

A. Vlasova, V. Nazarov, I. Yarova, N. Kushnirenko

National Odesa Polytechnic University
1, Shevchenko Ave., Odesa, 65044, Ukraine
Email: kushnirenko@op.edu.ua

The article presents a system for automated analysis of user digital footprint in social media using Open Source Intelligence (OSINT) approaches adapted for the Ukrainian language. Growing user activity on social networks creates privacy risks through accumulation of personal information. The aim of the work is to create a specialized system for comprehensive analysis of user digital presence in Telegram and YouTube considering morphological features of Ukrainian language and cultural context. Analysis of existing social platform research methods revealed limitations when processing Ukrainian content, particularly insufficient recognition accuracy and absence of specialized linguistic models. The proposed system consists of three sequential algorithms: a data collection algorithm through Telegram and YouTube APIs, a comprehensive text content analysis algorithm using TF-IDF method for key term extraction and langdetect library for automatic language detection, and an integrated profile construction algorithm with heterogeneous data normalization based on textual, social, temporal, and behavioral components. Experimental validation on a sample of 100 users across different age categories demonstrated system accuracy of 73% for Ukrainian content with average processing time of 85-95 seconds per user. The highest accuracy was achieved in determining temporal activity patterns (87%) and social activity level (77%). Scientific novelty lies in creating a specialized system for Ukrainian language considering its morphological features, including integration of TF-IDF frequency analysis methods adapted for word forms specificity, development of Ukrainian internet slang dictionary with over 800 entries, and construction of a comprehensive digital profile model based on four components.

Keywords: OSINT, digital footprint, social media, Telegram, YouTube, machine learning, text analysis.

DOI 10.15276/imms.v15.no4.545
УДК 004.056.5

Informatics and Mathematical Methods in Simulation
Vol.15 (2025), No. 4, pp. 545-553

УДОСКОНАЛЕННЯ СТЕГАНОГРАФІЧНОГО МЕТОДУ ВБУДОВУВАННЯ ІНФОРМАЦІЇ В ЧАСТОТНУ ОБЛАСТЬ ЦИФРОВИХ ЗОБРАЖЕНЬ

В. Ю. Волошин, В. В. Подуфалов, Г. Р. Пашнєв, Н. І. Кушніренко

Національний університет «Одеська політехніка»
1, Шевченка пр., Одеса, 65044, Україна
Email: 1945vlad1945@gmail.com

На сьогоднішній день існує значна кількість методів та інструментів для приховування інформації у цифрових зображеннях. Проте більшість із них залишаються вразливими до стиснення, фільтрації або інших атак, що призводить до спотворення або втрати прихованих даних. Особливо це стосується методів, які працюють у просторовій області, де навіть незначна обробка зображення може зруйнувати вбудоване повідомлення. Метою роботи є підвищення стійкості та надійності приховування інформації у цифрових зображеннях шляхом удосконалення стеганографічного методу Коха і Жао, який працює в частотній області з використанням дискретного косинусного перетворення (ДКП). У роботі проведено аналіз існуючих стеганографічних методів і програмних засобів, визначено їх переваги та недоліки, що дозволило сформувати напрямок подальшого удосконалення. Розроблений метод включає використання керованого вибору блоків ДКП за допомогою спеціального ключа, що забезпечує рівномірний розподіл прихованих даних і підвищує стійкість методу до JPEG-стиснення. Проведено експериментальні дослідження, які показали покращення показників пікового співвідношення сигналу до шуму (PSNR) у порівнянні з базовим алгоритмом, а також збереження коректності декодування даних навіть після стиснення зображення до 60%. Розроблений метод може бути застосований у системах захисту інформації, цифрового водяного маркування, а також у програмних рішеннях для приховування конфіденційних даних. Отримані результати підтверджують доцільність використання удосконаленого методу Коха і Жао для підвищення надійності та безпеки цифрових зображень. Використання ключового механізму вибору блоків відкриває можливості для побудови більш складних систем багаторівневого приховування інформації. Отримані результати створюють підґрунтя для подальших досліджень у напрямку комбінування стеганографічних і криптографічних методів для забезпечення комплексного захисту цифрових даних.

Ключові слова: стеганографія, метод Коха і Жао, частотна область, цифрове зображення.

Вступ. У сучасному цифровому світі обсяги обміну інформацією невпинно зростають, що підвищує вимоги до захисту даних від несанкціонованого доступу, підробки та перехоплення. Одним із ефективних напрямів забезпечення інформаційної безпеки є стеганографія — метод приховування фактів передавання даних у мультимедійних об’єктах, зокрема у цифрових зображеннях. На відміну від криптографії, стеганографія не лише зберігає конфіденційність повідомлення, але й приховує сам факт його існування, що робить її особливо корисною у сферах безпечного обміну інформацією, цифрового водяного маркування та захисту авторських прав. Попри значний прогрес у розробці стеганографічних методів, багато з них залишаються вразливими до сучасних атак — стиснення JPEG, фільтрації, зміни розміру або конвертації формату. Це особливо стосується методів, що працюють у просторовій області, де навіть незначне редагування зображення може призвести до часткової або повної втрати прихованих даних. Саме тому останніми роками активно розвиваються методи, які використовують частотну область, зокрема на основі дискретного косинусного перетворення. Одним із найбільш відомих і поширених методів частотної стеганографії є метод Коха і Жао, який базується на модифікації коефіцієнтів ДКП. Цей метод характеризується високою

ефективністю, але водночас має обмеження, пов'язані з рівномірністю вибору блоків для вбудовування інформації та зниженням якості зображення при підвищенні стійкості. Уразливість алгоритму до геометричних атак і до втрат при JPEG-стисненні зумовлює необхідність його вдосконалення. З огляду на це, метою даної роботи є удосконалення стеганографічного методу Коха і Жао шляхом запровадження ключового механізму вибору блоків ДКП, що забезпечує більш рівномірний розподіл даних, а також можливість керування параметром вбудовування для досягнення оптимального співвідношення між якістю зображення та стійкістю прихованої інформації. У процесі дослідження було проведено аналіз існуючих методів стеганографії, їх сильних і слабких сторін, а також експериментальне порівняння удосконаленої версії алгоритму з базовим методом. Результати показали, що нова модифікація дозволяє зберегти коректність відновлення даних навіть після суттєвого JPEG-стиснення (до 60%) і забезпечує вищі значення пікового співвідношення сигналу до шуму. Таким чином, розроблений підхід сприяє підвищенню надійності, стійкості та ефективності стеганографічних систем, що підтверджує його актуальність для сучасної кібербезпеки та практичного застосування у сфері захисту цифрової інформації [1].

Мета і задачі роботи. Метою роботи є підвищення стійкості та надійності приховування інформації у цифрових зображеннях шляхом удосконалення стеганографічного методу Коха і Жао з використанням механізму вибору блоків дискретного косинусного перетворення на основі ключа. Для досягнення поставленої мети необхідно розв'язати такі завдання:

- провести аналіз існуючих методів стеганографії, визначити їх переваги та недоліки, а також обґрунтувати вибір методу Коха і Жао як базового;
- дослідити особливості вбудовування інформації в частотну область за допомогою дискретного косинусного перетворення;
- розробити удосконалену модифікацію методу Коха і Жао із використанням ключового механізму вибору блоків для приховування даних;
- провести експериментальні дослідження ефективності запропонованого методу, порівняти його з базовим алгоритмом за показниками якості зображення та стійкості до JPEG-стиснення;

Основна частина. Методи просторової області в стеганографії охоплюють техніки, де приховання даних відбувається безпосередньо в пікселі зображення. Головна мета полягає в тому, щоб ефект наявності повідомлення був непомітним для спостерігача при перегляді зображення. Існують різні способи класифікації методів стеганографії (рис. 1) [2]:

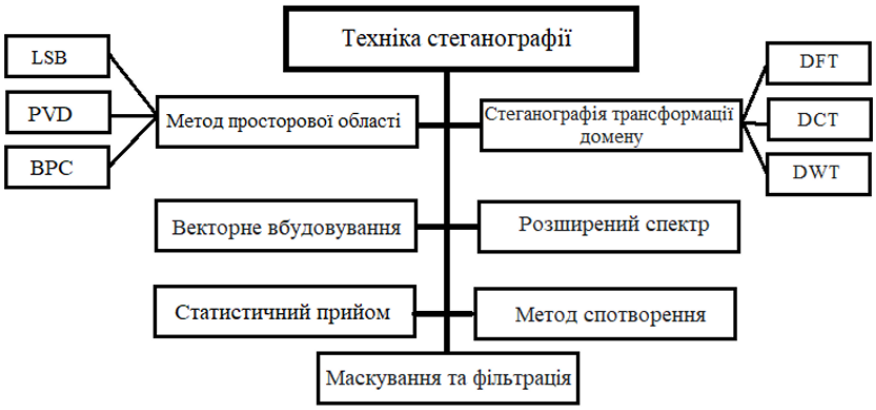


Рис.1. Техніка стеганографії

Дискретне косинусне перетворення (ДКП, DCT) є різновидом лінійного ортогонального перетворення, яке, на відміну від дискретного перетворення Фур'є, забезпечує більш ефективну енергетичну компресію, трансформуючи зображення з

просторової області у частотну, що дозволяє виділити значущі компоненти сигналу для подальшої обробки або стеганографічного вбудовування. Тобто представляє зображення у вигляді матриці 8×8, де зліва зверху знаходяться значення, що відповідають за фонові елементи зображення, а справа знизу – за контури (рис. 2) [3]. Процедура ДКП застосовується до кожного блоку від лівого верхнього кута до правого нижнього кута. Після цього кожен блок стискається з використанням таблиці квантування для масштабування коефіцієнтів ДКП, а потім у стислі блоки вбудовується повідомлення. При необхідності зображення може бути відновлене за допомогою процесу декомпресії, використовуючи зворотне дискретне косинусне перетворення [4].

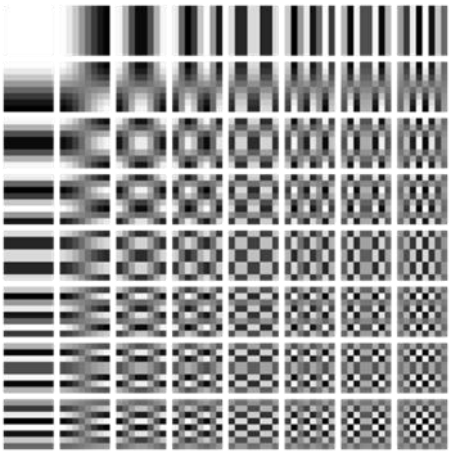


Рис.2. Блок частотної області ДКП

Дискретне косинусне перетворення використовується у JPEG стисненні. Процес стиснення включає в себе обнулення високочастотних складових матриці ДКП, а саме частот, які знаходяться в правому нижньому куті матриці. Ці високочастотні компоненти відповідають за різкі контури і деталі зображення. Тому під час стиснення JPEG саме на контурах можуть з'являтися артефакти, що проявляються у вигляді розмитих областей [5].

ДКП здійснюється за наступною формулою:

$$\mathrm{DCT}(i,j)=\frac{1}{\sqrt{2N}}C(i)C(j)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}pixel(x,y)\cos\left[\frac{(2x+1)i\pi}{2N}\right]\cos\left[\frac{(2y+1)j\pi}{2N}\right] \qquad (1.1)$$

де j — індекси коефіцієнта DCT у просторі частот;

y — координати пікселя у блоку зображення розміром $N\times N$;

$pixel(x, y)$ — значення інтенсивності пікселя у точці x,y (яскравість);

$C(i)$, $C(j)$ — нормувальні коефіцієнти [6].

Перейдемо до методу Коха і Жао, котрий будемо модифікувати. На початковому етапі первинне зображення стандартним чином розбивається на 8×8 блоки. До кожного блоку, який будемо позначати B, застосовується ДКП, тим самим здійснюючи переведення кожного блоку із просторової в частотну область. У результаті виходить 8×8 блок коефіцієнтів ДКП. Кожний блок призначено для приховання одного біта додаткової інформації (ДІ).

Існує дві реалізації алгоритму:

1. Для вбудови біта ДІ використовуються 2 коефіцієнта ДКП;
2. Для вбудови біта ДІ використовуються 3 коефіцієнта ДКП.

Розглянемо докладно перший варіант. Під час організації прихованого каналу зв'язку абоненти повинні попередньо домовитися (зв'язатися по захищеному каналу зв'язку) про два конкретних коефіцієнта ДКП із кожного блоку, які будуть використовуватися для приховання даних. Задамо дані коефіцієнти їх індексами $u1,v1$ і $u2,v2$ в масивах коефіцієнтів ДКП:

$$\begin{bmatrix} (1,1) & \cdots\cdots & (1,8) \\ \vdots & \cdots\ (u_1,v_1) & \cdots\ \vdots \\ \vdots & \cdots\ (u_2,v_2) & \cdots\ \vdots \\ (8,1) & \cdots\cdots & (8,8) \end{bmatrix}.$$

Відмітимо, що зазначені індекси повинні відповідати середньочастотним коефіцієнтам ДКП, що забезпечить: прихованість інформації; вбудована інформація не буде спотворюватися при Jpeg-стиску зі значними коефіцієнтами якості (або, що те ж саме, з малими коефіцієнтами стиску) [7].

На практиці найчастіше використовуються $u1,v1=4,5$ і $u2,v2=5,4$. Нехай у процесі стеганоперетворення треба вбудувати черговий біт $bk \in 0,1$ ДІ. Відповідно до секретного ключа для цього вибирається блок ЦЗ-контейнера. Відповідний йому блок коефіцієнтів ДКП позначимо $ВДК$:

$$B^{ДКП} = \begin{bmatrix} b_{11}^{ДКП} & b_{12}^{ДКП} & \ldots & b_{18}^{ДКП} \\ b_{21}^{ДКП} & b_{22}^{ДКП} & \ldots & b_{28}^{ДКП} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ b_{81}^{ДКП} & b_{82}^{ДКП} & \ldots & b_{88}^{ДКП} \end{bmatrix}.$$

Для вбудови bk використовуються коефіцієнти $bu1,v1ДКП,$. Вбудова біта bk відбувається таким чином: якщо $bk=0$, то різниця абсолютних значень використовуваних для вбудовування коефіцієнтів ДКП роблять більше деякої заданої додатної величини P, а якщо $bk=1$, то ця різниця робиться менше -P:

$$\begin{cases} \left|b_{u_1,v_1}^{ДКП}\right| - \left|b_{u_2,v_2}^{ДКП}\right| > P, \quad при \quad b_k = 0, \\ \left|b_{u_1,v_1}^{ДКП}\right| - \left|b_{u_2,v_2}^{ДКП}\right| < -P, \quad при \quad b_k = 1. \end{cases}$$

Таким чином, первинне зображення спотворюється за рахунок внесення змін у коефіцієнти ДКП, якщо їх відносні величини не відповідають приховуваному біту. Чим більше P, тим стеганосистема, створена на основі даного методу, є більш стійкою до стиску, однак якість зображення при цьому може значно погіршитися [8].

Після відповідного внесення корекції в значення коефіцієнтів ДКП, проводиться зворотне ДКП блоку. У результаті пересилання стеганоповідомлення, як вже зазначалося вище, зазнає спотворення, спотворення зазнає й ДІ. Для витягу ДІ виконується аналогічна процедура вибору коефіцієнтів ДКП у кожному блоці, що були задіяні в стеганоперетворенні, а розв'язок про переданий біт ухвалюється у відповідності з наступним правилом:

$$\begin{cases} b_k = 0, \quad при \quad \left|\bar{b}_{u_1,v_1}^{ДКП}\right| > \left|\bar{b}_{u_2,v_2}^{ДКП}\right|, \\ b_k = 1, \quad при \quad \left|\bar{b}_{u_1,v_1}^{ДКП}\right| < \left|\bar{b}_{u_2,v_2}^{ДКП}\right| \end{cases},$$

де $bu1,v1ДКП,$ - коефіцієнти ДКП блоку можливо зміненого при передачі стеганоповідомлення [9].

Пропонується додати ключ, за допомогою якого здійснюватиметься вибір блоків дискретного косинусного перетворення, у які буде приховуватися додаткова інформація. Використання ключа дозволить здійснити керований процес вибору блоків, що робить процедуру вбудовування більш організованою та послідовною.

На початковому етапі первинне зображення стандартним чином розбивається на 8×8 блоки однакового розміру. Розбиття виконується за тією ж схемою, що й в оригінальній модифікації методу, що дозволяє забезпечити узгодженість із базовим алгоритмом. Після цього до кожного з отриманих блоків застосовується дискретне косинусне перетворення. Завдяки цьому кожен блок зображення переводиться з просторової області в частотну. У результаті такого перетворення кожен блок зображення представляється у вигляді блоку коефіцієнтів ДКП, які описують його частотні характеристики. Це буде виглядати наступним чином (рис. 3) [10].

U_{11}	U_{12}	U_{13}	U_{14}	U_{15}	U_{16}	U_{17}	U_{18}
U_{21}	U_{22}	U_{23}	U_{24}	U_{25}	U_{26}	U_{27}	U_{28}
U_{31}	U_{32}	U_{33}	U_{34}	U_{35}	U_{36}	U_{37}	U_{38}
U_{41}	U_{42}	U_{43}	U_{44}	U_{45}	U_{46}	U_{47}	U_{48}
U_{51}	U_{52}	U_{53}	U_{54}	U_{55}	U_{56}	U_{57}	U_{58}
U_{61}	U_{62}	U_{63}	U_{64}	U_{65}	U_{66}	U_{67}	U_{68}
U_{71}	U_{72}	U_{73}	U_{74}	U_{75}	U_{76}	U_{77}	U_{78}
U_{81}	U_{82}	U_{83}	U_{84}	U_{85}	U_{86}	U_{87}	U_{88}

Рис.3. Блоки ДКП зі своєю нумерацією

У звичайному методі зазначається, що кожний блок призначений для приховування лише одного біта додаткової інформації. Такий підхід є доволі простим, але він не враховує можливість використання більш гнучкого механізму вибору блоків.

В адаптованому методі пропонується застосувати спеціальний ключ, який буде генеруватися самостійно. Основна вимога до ключа – відсутність повторюваних символів, що забезпечує унікальність та однозначність вибору блоків. Ключ формується у вигляді послідовності з п’яти двозначних чисел, які між собою не повторюються. Для прикладу візьмемо ключ «5312761837». Якщо розділити його на частини, отримаємо п’ять чисел: «53 12 76 18 37». У такому вигляді він зручніше сприймається та дозволяє наочно показати, яким чином відбувається відображення на блоки ДКП. Таким чином, кожне число відповідає конкретному блоку з певною нумерацією, і саме у ці блоки буде приховуватися додаткова інформація. На рисунку 4 подано приклад схематичного відображення вибору блоків відповідно до заданого ключа.

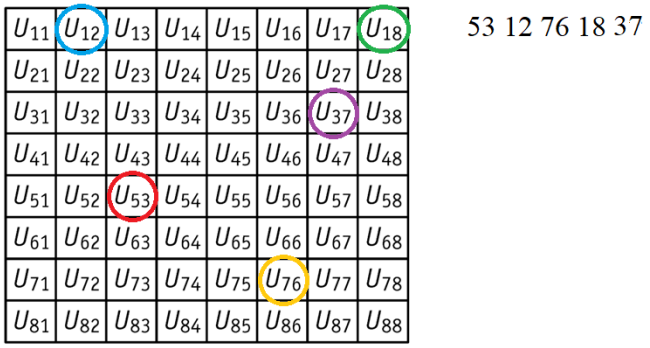


Рис.4. Вбудовування ДІ в ДКП блоки нумерація котрих відповідає ключу

Повторюємо дію для всіх чисел і отримуємо наступний вигляд ДКП блоків, якщо індекси блоків повторюються, то вони просто пропускаються, бачимо наступний рисунок (рис. 5).

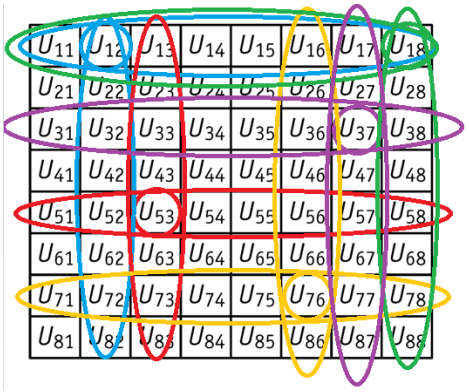


Рис.5. Кінцевий вид блоку в котрий будуть приховані дані

Тобто бачимо що блоки під індексами U21, U24, U25, U41, U44, U45, U61, U64, U65, U81, U84, U85 взагалі не будуть використанні на відміну від оригінального методу в котрому кожний блок призначено для приховання одного біта ДІ.

Далі все відбувається таким самим чином як і в оригінальній модифікації. Окрім того що буде додано можливість обирати значення P, для того щоб користувач сам міг обирати що йому більше потрібно, більший захист чи більша якість.

Експериментальні дослідження проводилися із використанням набору з 100 оригінальних зображень, які мали різний зміст та структуру. У кожне зображення вбудовувалася певна кількість додаткової інформації, виражена у відсотковому співвідношенні до максимально можливої кількості даних, які потенційно можуть бути розміщені у зображенні. Такий підхід дозволив оцінити, як змінюється якість зображень залежно від обсягу прихованої інформації.

У процесі експерименту було проведено серію тестів, під час яких дані вбудовувалися у синій канал зображення. При цьому параметр P було зафіксовано на рівні P = 25. Це дало змогу забезпечити однакові умови для всіх зображень та коректно порівняти результати.

Для оцінювання якості стеганографічних зображень було використано показник пікового співвідношення сигналу до шуму. Отримані значення PSNR дозволяють визначити ступінь спотворення зображення після процесу вбудовування інформації. Чим вищим є значення цього показника, тим менш помітними є зміни для людського ока.

На рисунку 6 зображені результати. Видно, що удосконалена модифікація дає кращі результати від вже наявної.

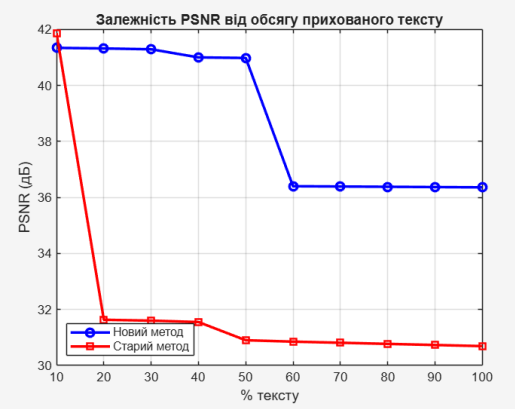


Рис.6. Графік залежності кількості тексту і впливу на PSNR зображення

Для більш ґрунтовної перевірки його можливостей було проведено додатковий експеримент, спрямований на оцінку здатності алгоритму коректно декодувати додаткову інформацію за умов різних рівнів стиснення JPEG.

У даному дослідженні аналізувалася робота методу при змінних значеннях коефіцієнта якості QF, а також при різних параметрах P. Це дозволило оцінити не лише загальну стійкість алгоритму, але й визначити, як саме змінюється його надійність у залежності від умов стиснення.

На рисунку 7 зображені результати. Отримані експериментальні результати підтверджують, що запропонована модифікація демонструє покращені характеристики порівняно з наявним базовим методом. Результати декодування ДІ при значеннях параметра P = 25 та P = 40 не демонструють помітних змін. Однак для значенням P = 55 стало можливим успішно витягти ДІ при стисненні 80%. Подальші експерименти показали, що при значеннях P = 70 та P = 85 цифрову інформацію вдалося відновити за умов стиснення до 70%. Найкращий результат було отримано при P = 100 — навіть за умови стиснення зображення до 60% цифрова інформація була витягнута без спотворень.

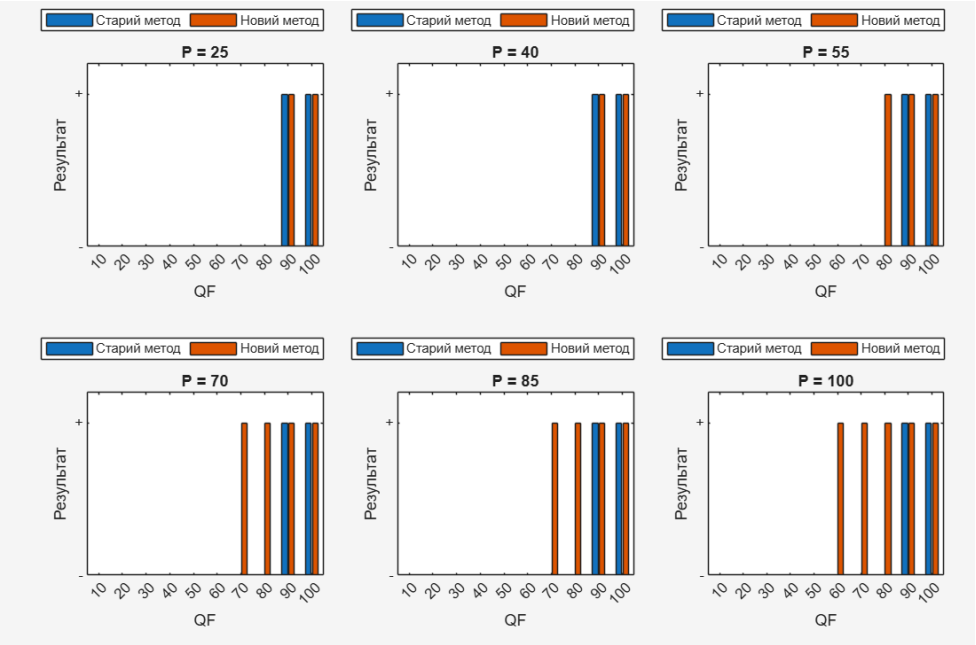


Рис.7. Гістограми можливості декодування ДІ в порівнянні двох методів

Висновки. Проведено дослідження методів стеганографії та їх практичного застосування для приховування інформації у цифрових зображеннях. Було виконано порівняльний аналіз сучасних стеганографічних технік, серед яких особливу увагу приділено методам, що працюють в частотній області. Це дозволило виявити основні недоліки традиційних підходів, зокрема недостатню стійкість до JPEG-стиснення та геометричних атак, що стало підґрунтям для подальшого удосконалення алгоритмів.

У роботі розроблено удосконалену модифікацію методу Коха і Жао, яка базується на використанні ключового механізму вибору блоків дискретного косинусного перетворення. Такий підхід забезпечує рівномірніший розподіл прихованої інформації, зменшує ймовірність спотворень та підвищує стійкість методу до втрат під час стиснення зображень.

Проведено експериментальні дослідження із використанням набору тестових зображень різного типу. За результатами експериментів підтверджено, що удосконалена модифікація демонструє вищі показники пікового співвідношення сигналу до шуму порівняно з базовим методом, а також зберігає можливість коректного декодування додаткової інформації навіть після стиснення зображень до 60% якості.

Розроблений метод відзначається високим рівнем надійності, стійкості та непомітності, що дозволяє рекомендувати його для практичного використання у системах захисту інформації, цифрового водяного маркування та інших задач кібербезпеки, де важливим є збереження якості зображення та безпека переданих даних.

Отже, результати даного дослідження мають як теоретичне, так і практичне значення, оскільки демонструють можливість ефективного удосконалення класичних стеганографічних методів шляхом введення додаткових керованих параметрів та адаптивного вибору блоків в частотній області. Подальші дослідження можуть бути спрямовані на розширення функціональності методу, його інтеграцію з криптографічними механізмами та застосування в системах багаторівневого захисту цифрових медіа.

Список літератури

1. Кулик М.В. Дослідження сучасних алгоритмів побудови цифрових водяних знаків для відео-контенту. Київ: Київський політехнічний інститут імені Ігоря Сікорського. 2018. 40 с.
2. Laskar B., Bouzid M. Enhancing secure communication: a QIM-based steganography approach for G. 722.2 speech streams with Stable Roommate Index Division. *Multimedia Tools and Applications*. 2024. P. 1-19.
3. Agarwal S., Jung K.H. Digital image steganalysis using entropy driven deep neural network. *Journal of Information Security and Applications*. 2024. V. 84. P.103799.
4. Zhang C., Jiang S., Chen Z. SPM: estimating payload locations of QIM-based steganography in low-bit-rate compressed speeches. *Multimedia Tools and Applications*. 2024. P. 1-26.
5. Cohen R. Cryptanalysis of Practical Optical Layer Security Based on Phase Masking of Mode-Locked Lasers and Multi-Homodyne Coherent Detection. *Journal of Lightwave Technology*. 2024.
6. Cohen R. Cryptanalysis of Practical Optical Layer Security Based on Phase Masking of Mode-Locked Lasers. *Journal of Lightwave Technology*. 2023.
7. Нищик В.І. Розробка мобільного додатка для Android з реалізацією методу LSB для стеганографії. Одеса: Одеський державний екологічний університет, 2022. 12 с.
8. Discrete cosine transform. URL: https://www.mathworks.com/help/signal/ref/dct.html
9. Зоріло В.В., Лебедєва О. Ю., Петрук К. О. Виявлення мультиплікативного шуму в цифрових зображеннях в умовах збереження з втратами. Одеса: НУОП, 2023.
10. Dixit M., Bhide N., Khankhoje S., Ukarande R. Video Steganography. *Pervasive Comput*. 2021. V. 1. P. 1–4.

IMPROVEMENT OF A STEGANOGRAPHIC METHOD FOR EMBEDDING INFORMATION INTO THE FREQUENCY DOMAIN OF DIGITAL IMAGES

V. Voloshyn, V. Podufalov, H. Pashniev, N. Kushnirenko

National Odesa Polytechnic University
1, Shevchenko Ave., Odesa, 65044, Ukraine
Email: 1945vlad1945@gmail.com

Today, there are a significant number of methods and tools for hiding information in digital images. However, most of them remain vulnerable to compression, filtering, or other attacks, leading to distortion or loss of hidden data. This is especially true for methods that operate in the spatial domain, where even minor image processing can destroy the embedded message. The aim of this work is to improve the robustness and reliability of information hiding in digital images by improving the steganographic method of Koch and Zhao, which works in the frequency domain using discrete cosine transform (DCT). The paper analyses existing steganographic methods and software tools, identifies their advantages and disadvantages, and suggests directions for further improvement. The developed method involves the use of controlled selection of DCT blocks using a special key, which ensures uniform distribution of hidden data and increases the stability of the method to JPEG compression. Experimental studies have been conducted, which showed an improvement in peak signal-to-noise ratio (PSNR) compared to the baseline algorithm, as well as the preservation of data decoding accuracy even after image compression to 60%. The developed method can be applied in information security systems, digital watermarking, and software solutions for hiding confidential data. The results obtained confirm the feasibility of using the improved Koch and Zhao method to improve the reliability and security of digital images. The use of a key block selection mechanism opens up opportunities for building more complex multi-level information hiding systems. The results obtained provide a basis for further research into combining steganographic and cryptographic methods to ensure comprehensive protection of digital data.
Keywords: steganography, Koch and Zhao method, frequency domain, digital image.

DOI 10.15276/imms.v15.no4.554
УДК 004.056.5

Informatics and Mathematical Methods in Simulation
Vol.15 (2025), No. 4, pp. 554-566

ВИКОРИСТАННЯ ТЕОРІЇ ГРАФІВ В УМОВАХ СЕЛЕКТИВНОЇ СТЕГАНОГРАФІЇ

С. М. Григоренко[1], А. А. Кобозєва[2]

[1]Національний університет «Одеська політехніка»
1, Шевченка пр., Одеса, 65044, Україна
[2]Одеський національний морський університет
34, Мечникова вул., Одеса, 65029, Україна
Email: alla_kobozeva@ukr.net

Стеганографія є сьогодні одним з найпотужніших засобів захисту інформації, проте велика кількість сучасних стеганометодів мають обмеження області застосування, не розраховані на ефективну роботу з випадковим контейнером, який на сьогодні є найпоширенішим при практичному використанні стеганосистеми. Для рішення цієї проблеми використовується селективна стеганографія, яка значно покращує характеристики стеганосистеми, зокрема дозволяє зменшити спотворення контейнеру, в якості якого в роботі розглядається цифрове зображення, в результаті стеганоперетворення. На сьогодні проблеми селективної стеганографії не є достатньо дослідженими, відсутні загальні принципи селекції, які б не були розраховані на конкретні стеганоалгоритми, практично не розглядаються питання стійкості результату селекції до збурних дій. Метою даної роботи є удосконалення селективного методу, запропонованого авторами раніше, для підвищення стійкості результату селекції блоків до атак проти вбудованого повідомлення. Мета досягається шляхом побудови бінарного відношення нестрогого порядку, визначеного на множині X непересічних блоків матриці контейнера, отриманих шляхом її стандартної розбивки, граф якого використовується для організації селекції блоків, при цьому кількісним критерієм для вибору блоку B є нормована відокремленість максимального сингулярного числа блоку BB^T. Найбільш важливим результатом роботи є розробка та застосування в запропонованому методі загальних принципів селекції блоків, що не є орієнтованими на конкретні стеганоалгоритми. Практична значимість отриманих результатів полягає в підвищенні стійкості результату селекції в умовах збурних дій, що забезпечується запропонованим методом, наслідком чого є підвищення ефективності стеганографічної системи в цілому.

Ключові слова: селективна стеганографія, бінарне відношення порядку, граф бінарного відношення, стійкість до збурних дій.

Вступ. Проблеми інформаційної безпеки та захисту інформації сьогодні є вкрай актуальними для України в умовах повномасштабної війни, яку розв'язала російська федерація, що включає, як складову, інформаційну агресію проти українського суспільства. Безпрецедентне зростання кібератак, інформаційно-психологічних операцій та цифрового шпигунства вимагає від держави, військових, критичної інфраструктури застосування найсучасніших та найнадійніших методів приховування та передачі даних.

Одним з найпотужніших засобів захисту інформації в умовах постійного зовнішнього моніторингу та аналізу трафіку є сьогодні стеганографія [1,2] – наука і мистецтво приховування інформації. На відміну від криптографії, результат застосування якої зразу «кидається в очі», представляючи «нечитабельний» контент, стеганографія приховує від третіх осіб сам факт передачі секретного повідомлення, що критично важливо для уникнення виявлення та протидії з боку супротивника. Проте велика кількість сучасних стеганометодів мають обмеження області застосування, вони не розраховані на ефективну роботу з випадковим контейнером, в якості якого далі розглядається цифрове зображення (ЦЗ), оскільки передбачають використання для

вбудови додаткової інформації (ДІ) всього контейнера чи його частин, зокрема блоків, обраних випадковим чином. Для деяких стеганометодів це несистематично призводить до невиконання певних вимог, що висуваються до відповідної стеганосистеми, при їх використанні, зокрема до порушення надійності сприйняття стеганоповідомлення, тобто до візуально помітної відмінності ЦЗ-контейнера від зображення, що несе в собі ДІ. Але саме випадковий контейнер є найпоширенішим для сучасних стеганосистем. Для покращення результату його використання довільним стеганометодом все більшого значення набуває селективна стеганографія, метою якої є обрання таких ділянок контейнера, які є найбільш сприятливими в тому чи іншому сенсі для стеганоперетворення, зокрема таких, що забезпечують відносно більший ступінь «непомітності» змін ЦЗ при вбудові ДІ в порівнянні з випадковим/довільним вибором ділянок контейнера для стеганоперетворення. Завдяки вибірковому підходу, селективна стеганографія дозволяє мінімізувати перцептивні спотворення на стеганоповідомленні, що є важливим для уникнення підозр з боку порушників.

Таким чином, дослідження, розробка та удосконалення методів селективної стеганографії є сьогодні критично актуальним завданням для забезпечення надійного та непомітного каналу передачі даних в умовах зовнішньої агресії та постійного кібернагляду.

Стан проблеми. На сьогоднішній день проблеми селективної стеганографії не є достатньо дослідженими: відсутні загальні принципи селекції, які б не були орієнтовані на конкретні стеганоалгоритми; часто селекція відбувається в просторовій області ЦЗ-контейнера, де визначальними критеріями відбору стають певні піксельні характеристики, а результатом – набір пікселів, хоча областю вбудови ДІ може слугувати область перетворення контейнера (частотна, області сингулярного, спектрального розкладання тощо); практично не розглядаються питання стійкості результату селекції до збурних дій.

Серед вимог до сучасної стеганографічної системи одною з основних є вимога забезпечення надійності сприйняття формованого стеганоповідомлення, в світлі чого і розглядатимуться можливості і переваги селективної стеганографії.

На сьогоднішній день найчастіше селекція областей ЦЗ-контейнера для вбудови ДІ зводиться до пошуку в ньому контурів [3-6]. Це є абсолютно природним, оскільки особливістю зорової системи людини є набагато більша ймовірність помітити зміни на зображенні в областях з малими перепадами значень яскравості, ніж в областях, де наявні контури, маленькі деталі, тобто перепади значень яскравості є значними. Підходи, що застосовуються в відповідних роботах, є різними: комплексне вейвлет-перетворення з подвійним деревом [3], де для стеганоперетворення використовуються коефіцієнти дискретного вейвлет-перетворення, а високочастотні області зображення ідентифікуються за допомогою методу виявлення контурів Кенні; детектор контурів Кірша [5]; в [4] використовується глибоке навчання, за допомогою якого отримується мапа контурів у вигляді двійкового зображення; в [6] контури виявляються за допомогою фільтра Собеля.

Хоча, з одного боку, виявлення контурів ЦЗ як результату селекції, як зазначено вище, є природним з урахуванням зазначеної мети відбору, з іншого боку, бажаним результатом тут є забезпечення покращення якісного/кількісного стану стеганоповідомлення для будь-якого зображення-контейнера, навіть, якщо це зображення взагалі не містить об'єктів/контурів, як, наприклад, отримане непрофесійною відеокамерою, представлене на рис.1. Для таких зображень орієнтація на контури, як області для вбудови ДІ, призведе до катастрофічного падіння (аж до 0 біт/піксель) пропускної спроможності прихованого каналу зв'язку, що не дасть можливості використовувати їх в якості контейнерів, а тому такий селективний підхід принципово не може забезпечити можливість використання випадкового контейнера довільним стеганометодом. Крім того, цей підхід є досить «категоричним», однозначно

локалізуючи області вбудови ДІ в межах контурів, тим самим чітко обмежуючи зверху обсяг ДІ, яка може бути вбудована в контейнер.

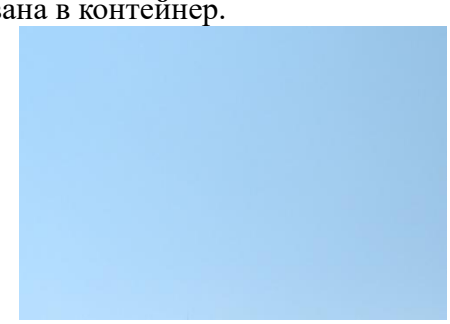


Рис. 1. Оригінальне ЦЗ, що не містить контурів

Велика кількість сучасних селективних методів, зокрема і ті, що розглянуті вище, розроблені під метод модифікації найменшого значущого біта (LSB-метод) [7], що, на погляд авторів даної роботи, знижує цінність результатів, що представляються, оскільки, по-перше, це звужує область застосування селективного методу (чи вимагає додаткових досліджень для обгрунтування можливості/неможливості його використання під інші стеганометоди), по-друге, передбачає проведення стеганоперетворення лише в просторовій області контейнера, виключаючи стеганометоди, що працюють з областями перетворення, по-третє, LSB-метод сам по собі є таким, що забезпечує надійність сприйняття формованного стеганоповідомлення без будь-якої селекції.

Зауважимо, що, крім вже визначених недоліків, розглянуті вище роботи є показовими з точки зору існуючої загальної тенденції «ігнорування» складної проблеми, що залишається недостатньо дослідженою, хоча є критично важливою для організації ефективного декодування ДІ в умовах атак проти вбудованого повідомлення, тобто процесу, що є ключовим при організації прихованого каналу зв'язку, в умовах, що мають високу ймовірність: забезпечення стійкості результату проведеної селекції до збурних дій. Це питання майже не розглядається в межах селективної стеганографії.

Попіксельна селекція, представлена в роботах, згаданих вище, не завжди є виправданою не тільки по причині орієнтованості лише на просторове стеганоперетворення. На сьогодні найпоширенішими стеганографічними методами є блокові, де перед вбудовою ДІ матриця контейнера піддається розбивці на блоки, а селекція природно зводиться до вибору найбільш сприятливих для стеганоперетворення блоків. Один з найбільш ефективних блокових селективних методів, запропонований в [8], забезпечує, як заявляють автори, для відповідного стеганоповідомлення значення показника пікового відношення «сигнал-шум» *PSNR* 61-65 dB навіть в умовах 80% модифікованих блоків. Запропонована схема розглядає виключно просторову область контейнера для стеганоперетворення, обираючи певний «регіон», при цьому цей вибір робиться серед пікселів в межах блоку. Метод вбудовує три біти ДІ в крайових пікселях з використанням мінімальної середньоквадратичної похибки для визначення можливості/неможливості залучення пікселів внутрі блоку. Таким чином, авторами пропонується селективний метод під конкретний спосіб безпосередньої вбудови ДІ, не даючи уявлення про якість селекції у випадку застосування іншого методу стеганоперетворення.

Загальний селективний блоковий метод, тобто такий, який ніяк не орієнтований на використовуваний для стеганоперетворення стеганоалгоритм, був нещодавно представлений в [9]. Вибір відбувається серед блоків, отриманих шляхом стандартної розбивки [10] матриці ЦЗ, з метою збільшення ступеня надійності сприйняття стеганоповідомлення, що кількісно оцінюється за допомогою різницевого показника візуального спотворення *PSNR* [11], в порівнянні з випадковим використанням блоків. Основною кількісною характеристикою блоку B – критерієм, що використовувався для

вибору, була обґрунтовано обрана нормована відокремленість його максимального сингулярного числа. Даний метод дозволив значно збільшити показник *PSNR* для отримуваних стеганоповідомлень незалежно від використовуваного стеганометоду. Хоча в [9] питання забезпечення стійкості результату проведеної селекції до атак проти вбудованого повідомлення – збурних дій не розглядалося, враховуючи загальність і ефективність метода [9], доцільним є, зберігаючи його теоретичні основи, віднайти шляхи для вирішення означеного питання.

Метою роботи є удосконалення селективного методу, запропонованого в [9], що забезпечить підвищення стійкості результату селекції блоків до атак проти вбудованого повідомлення.

Для досягнення мети в роботі вирішуються наступні задачі:

1. Зміна принципу побудови бінарного відношення на множині блоків ЦЗ-контейнера відносно застосовуваного в [9] для уникнення негативних наслідків особливостей машинної арифметики та забезпечення повного впорядкування множини блоків;
2. Зміна критерію вибору блоків для підвищення стійкості селекції до атак проти вбудованого повідомлення;
3. Розробка удосконаленого селективного методу;
4. Оцінка ефективності запропонованого селективного методу.

Вступні зауваження. Для кращого розуміння шляхів удосконалення селективного методу [9] розглянемо детально спосіб організації вибору блоків, що ним реалізується.

Нормована відокремленість максимального сингулярного числа будь-якого блоку B $svdgap_n(1,B)$, що є критерієм, який використовувався для вибору блоків в [9], визначається у відповідності з формулою:

$$svdgap_n(1,B) = \bar{\sigma}_1 - \bar{\sigma}_2, \qquad (1)$$

де $\bar{\sigma}_i = \dfrac{\sigma_i}{\|\sigma\|}$ – нормовані сингулярні числа (СНЧ), $\sigma = (\sigma_1(B), \sigma_2(B), ..., \sigma_l(B))$ – вектор СНЧ B, $\|\cdot\|$ – Евклідова векторна норма,

Нормований вектор СНЧ

$$\bar{\sigma} = \left(\bar{\sigma}_1(B), \bar{\sigma}_2(B), ..., \bar{\sigma}_l(B)\right) \qquad (2)$$

Для організації селекції на множині X блоків ЦЗ-контейнера будувалося бінарне відношення ρ наступним чином: впорядкована пара $< B_{ij}, B_{km} > \in \rho$, де $B_{ij}, B_{km} \in X$, якщо показники вкладу високочастотної складової, в якості яких застосовуються значення нормованої відокремленості максимальних СНЧ B_{ij}, B_{km}, будуть рівні:

$$B_{ij} \rho B_{km} \quad \Leftrightarrow \quad svdgap_n(1, B_{ij}) = svdgap_n(1, B_{km}). \qquad (3)$$

Отримане бінарне відношення є відношенням еквівалентності, тому розбиває множину X на непересічні класи еквівалентності, в кожному з яких опиняються блоки контейнера, пов'язані відношенням ρ, а для будь-якої пари блоків з різних класів зв'язок їх в сенсі ρ є відсутнім. Відношенню ρ ставився у відповідність орієнтований граф, кожна компонента зв'язності якого визначала певний клас. За допомогою простої гомоморфної згортки вершин графа, що знаходилися в одній компоненті зв'язності, будувався зважений макрограф, що відповідав ЦЗ-контейнеру, з макровершинами, що визначали класи еквівалентності блоків зображення. Вага макровершини відповідала значенню нормованої відокремленості максимального СНЧ блоків цього класу. Для вибору безпосередніх блоків для вбудови ДІ, що розглядалася як бінарна послідовність $p_1, p_2, ..., p_t$, $p_i \in \{0,1\}$, визначалась загальна кількість T блоків контейнера, необхідних для вбудови наявної ДІ використовуваним стеганоалгоритмом. Для стеганоперетворення використовувалися T блоків, що відповідали макровершинам

побудованого макрографа контейнера, починаючи з макровершини з найменшою вагою (вклад високо-, середньочастотної складової в цих блоках є найбільшим, а тому ймовірність виникнення артефактів після вбудови ДІ є порівняно малою) в порядку зростання ваги вершин. Блоки з одного класу еквівалентності обиралися випадковим чином.

Принцип побудови бінарного відношення (3) не позбавлений недоліку: встановлення рівності для дійсних чисел $svdgap_n(1, B_{ij})$, $svdgap_n(1, B_{km})$, що є результатами обчислень в системі чисел з плаваючою точкою, є нетривіальною задачею в силу скінченності множини таких чисел і, як наслідок, наявності обчислювальної похибки. І хоча всі СНЧ є добре обумовленими [12], що призводить до доброї обумовленості і нормованих відокремленостей максимальних СНЧ блоків, два природно рівних (нерівних) дійсних числа $svdgap_n(1, B_{ij})$, $svdgap_n(1, B_{km})$ в результаті обчислень можуть виявитися нерівними (рівними) в силу особливостей машинної арифметики. Для $svdgap_n(1, B_{ij})$, $svdgap_n(1, B_{km})$ по ходу організації процесу селекції у відповідності з [9] проводилось округлення до певної кількості значущих цифр, що є додатковим джерелом похибки в значеннях нормованої відокремленості максимального СНЧ блоку. Все вищезазначене об'єктивно може призвести до того, що блок контейнера опиниться не в тому класі еквівалентності, де він повинен бути природно, наслідком чого стануть похибки при побудові графа і відповідно макрокрафа ЦЗ-контейнера, знижуючи ефективність селекції.

Принцип побудови бінарного відношення на множині блоків ЦЗ-контейнера. Пропонується зміна принципа (3) побудови бінарного відношення ρ на множині X блоків ЦЗ-контейнера на наступний:

$$B_{ij}\rho B_{km} \quad \Leftrightarrow \quad svdgap_n(1, B_{ij}) \geq svdgap_n(1, B_{km}). \qquad (4)$$

Можливість рівності для $svdgap_n(1, B_{ij})$, $svdgap_n(1, B_{km})$ для забезпечення $B_{ij}\rho B_{km}$ в (4) залишається, що дозволяє сподіватися на збереженні всіх позитивних якостей селекції [9], але вона перестає бути єдиною визначальною властивістю.

Визначальна властивість (4) є менш чутливою до обчислювальної похибки, ніж (3), тому запропонована заміна сприятиме підвищенню стійкості результату селекції блоків в цілому до збурних дій, що підтверджується нижче результатами обчислювального експерименту.

Оскільки бінарне відношення (4) є рефлексивним, антисиметричним і транзитивним, воно є відношенням нестрогого порядку. При цьому будь-які два блоки з множини X є порівняними по відношенню порядку ρ, що призведе до того, що відношення ρ визначає повний порядок на множині X.

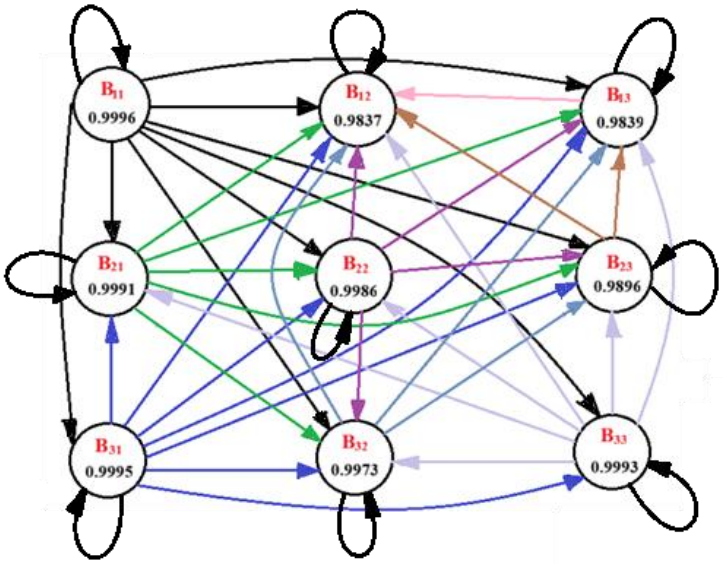
Бінарному відношенню (4) поставимо у відповідність орієнтований зважений граф $G(X, E)$, де X – множина вершин, E – множина ребер, вершини якого відповідають блокам ЦЗ, а вага вершини дорівнює нормованій відокремленості його максимального СНЧ (1). Для наочності на рис.2 наведено приклад побудови $G(X, E)$ для незначної за розміром частини ЦЗ (на рис.2(а) виділена червоним квадратом). Ваги вершин вказані всередині вершин.

Оскільки граф $G(X, E)$ є орієнтованим, для кожної вершини, можна визначити два локальних ступеня: $\rho^+(B_{ij})$ - кількість ребер графа, інцидентних вершині B_{ij}, для яких вершина B_{ij} є початком, $\rho^-(B_{ij})$ - кількість інцидентних вершині B_{ij} ребер, для яких B_{ij} є кінцем. Після того, як граф $G(X, E)$ побудований, для селекції блоків сам кількісний критерій (1) вже непотрібний, оскільки вибір блоків пропонується реалізовувати з урахування лише $\rho^+(B_{ij})$ і $\rho^-(B_{ij})$ для кожної вершини.

а

60	60	58	58	62	65	70	70	68	71	59	48	68	95	131	161	146	123	140	155	156	153	152	152
59	59	59	59	60	63	62	66	66	66	68	64	60	64	81	128	160	147	129	149	171	162	155	156
59	60	60	61	61	63	63	64	64	61	64	65	65	61	56	66	113	148	143	149	160	164	159	160
59	59	60	60	B11	60	61	61	61	61	62	62	B12	67	61	52	55	91	116	136	B13	160	156	163
58	58	57	57		58	59	60	60	60	62	61		67	65	65	51	48	74	105		157	164	159
58	56	53	52	54	56	57	59	57	59	59	60	62	63	62	64	70	63	54	62	83	128	164	169
57	54	54	53	53	54	56	57	56	59	60	59	59	59	60	57	67	86	76	65	59	73	121	159
56	55	56	54	53	54	54	55	54	60	60	56	53	57	58	58	61	71	79	76	78	56	63	106
56	57	55	54	52	53	55	53	54	57	57	52	54	56	57	58	59	60	68	72	75	78	62	56
55	55	55	54	53	55	57	55	53	57	58	58	57	57	57	57	57	59	64	66	65	73	81	80
55	53	56	57	56	53	56	60	57	57	59	59	58	57	57	57	55	54	59	62	62	66	74	83
56	57	58	58	57	55	57	62	68	61	58	59	59	59	58	57	52	55	59	60	62	61	62	65
57	58	59	59	B21	56	59	59	59	59	59	61	B22	70	60	58	58	57	58	59	B23	58	61	62
53	55	57	58	61	60	61	61	58	59	55	59	67	60	54	58	65	60	57	56	56	57	58	60
51	53	54	57	62	62	63	64	60	55	53	56	58	59	55	60	83	93	67	55	57	57	57	57
48	51	53	55	59	61	62	64	63	58	55	57	58	59	59	74	101	96	73	62	57	57	57	56
49	50	53	54	57	58	62	63	63	58	55	56	57	58	63	69	74	72	68	62	59	58	56	55
51	53	54	54	55	58	60	65	65	61	54	55	56	57	62	65	62	65	63	60	60	58	57	56
54	55	56	57	59	59	61	62	69	71	62	56	54	58	61	62	60	57	58	57	57	56	55	56
55	56	57	58	B31	60	59	62	67	71	73	62	B32	59	59	57	54	53	55	56	B33	55	55	55
52	55	56	59		61	60	60	60	61	64	65		61	59	54	55	55	55	55		54	54	55
53	55	58	62	62	62	61	59	55	58	60	61	62	63	60	56	57	56	56	55	55	55	54	53
55	54	58	65	64	61	61	60	56	53	57	58	62	69	64	60	60	58	55	55	56	55	53	52
57	55	58	60	60	60	60	58	56	56	57	55	58	70	65	61	61	59	58	58	57	55	52	51

б



в

Рис. 2. Приклад побудови графа $G(X,E)$ для ЦЗ: а – оригінальне ЦЗ, частина якого розміром 24×24 пікселя використовується для побудови графа; б – матриця обраної частини ЦЗ з відповідною стандартною розбивкою на блоки; в – граф $G(X,E)$

$G(X,E)$.

При цьому схема збереження побудованого графа ЦЗ для організації такої селекції може бути організована в спрощеному вигляді, зберігаючи лише інформацію про ступені вершин. Крім того, саме повна впорядкованість множини X, що досягається завдяки введеному бінарному відношенню порядку, дозволяє провести додаткові спрощення для схеми збереження графа

Будь-які блоки ЦЗ є порівнянними по введеному бінарному відношенню ρ, с чого випливає, що загальна кількість ребер, інцидентних кожній вершині, буде дорівнювати N-1, без урахування петель, де N – загальна кількість блоків матриці ЦЗ. Наявні петлі, що відповідають кожній вершині графа, говорячи про рефлексивність бінарного

відношення, для проведення селекції блоків не відіграють ніякої ролі, оскільки відтворюють лише зв'язок блоку з самим собою, тому ці ребра далі не має сенсу враховувати, оскільки вони тільки ускладнюють загальну схему графа. Таким чином:

$$\rho^{+}\left(B_{ij}\right)+\ \rho^{-}\left(B_{ij}\right)=N-1. \qquad (5)$$

Враховуючи (5) при збереженні необхідної для селекції інформації про отриманий граф $G(X,E)$ ЦЗ можна зберігати лише інформацію або про $\rho^{+}\left(B_{ij}\right)$, або про $\rho^{-}\left(B_{ij}\right)$. Таким чином, побудований на рис.2 граф може зберігатися у вигляді:

$$\begin{pmatrix} 8 & 0 & 1 \\ 5 & 4 & 2 \\ 7 & 3 & 6 \end{pmatrix} \quad \text{або} \quad \begin{pmatrix} 0 & 8 & 7 \\ 3 & 4 & 6 \\ 1 & 5 & 2 \end{pmatrix},$$

для $\rho^{+}\left(B_{ij}\right)$, $\rho^{-}\left(B_{ij}\right)$ відповідно. Найбільший пріоритет для вбудови ДІ мають блоки, яким відповідають найменші значення $\rho^{+}\left(B_{ij}\right)$ в першій матриці, найбільші значення $\rho^{-}\left(B_{ij}\right)$ в другій.

Критерій вибору блоків для підвищення стійкості селекції до атак проти вбудованого повідомлення. Для селективної стеганографії важливим є стійкість відповідного алгоритму до збурних дій – атак: селективний алгоритм повинен бути таким, щоб навіть в умовах атаки зберегти можливість відновлення блокового стеганошляху під час декодування ДІ. Оскільки кількісним критерієм для вибору блоку є нормована відокремленість його максимального СНЧ (1), яка, як вже було зазначено вище, є добре обумовленою, то її зміни при збуреннях блоків будуть незначними. Але самі по собі значення нормованої відокремленості при обраному способі селекції не відіграють ключової ролі після побудови $G(X,E)$: для забезпечення можливості ефективного декодування ДІ після атак ключовим буде не абсолютна зміна значень (1), а відносне співвідношення між значеннями цього параметру у блоків. Таким чином, для забезпечення можливості ефективного відновлення ДІ граф ЦЗ $G(X,E)$ повинен залишитися без змін, в порівнянні з його первісним видом, при цьому зміни нормованої відокремленості можуть відбуватися. Ця ідея є основною для побудови селективного блокового методу, що пропонується в роботі. Вона і надалі буде використовуватися авторами як загальна ідея ефективної селекції: після забезпечення стійкості нової «похідної» структури, що відповідає ЦЗ і несе інформаціє про його блокові зв'язки у відповідності з обраним критерієм (в даному випадку похідна структура – це $G(X,E)$), для якої основними є її якісні характеристики, безпосередній кількісний критерій стає непотрібним, його можна відкинути. Але ключовим моментом тут є попереднє забезпечення стійкості похідної структури.

В [13,14] доведено, що нормований вектор СНЧ $\overline{\sigma}$ (2), що відповідає довільному $l\times l-$блоку B, будучи нечутливим, має більшу чутливість до збурних дій, ніж також нечутливий нормований вектор

$$\overline{\overline{\sigma}}=\left(\overline{\overline{\sigma}}_1\left(BB^T\right),\overline{\overline{\sigma}}_2\left(BB^T\right),...,\overline{\overline{\sigma}}_l\left(BB^T\right)\right)=\frac{\sigma_{BB^T}}{\left\|\sigma_{BB^T}\right\|}, \qquad (6)$$

де $\sigma_{BB^T}=\left(\sigma_1\left(BB^T\right),\sigma_2\left(BB^T\right),...,\sigma_l\left(BB^T\right)\right)$ - вектор СНЧ BB^T. Враховуючи це, для підвищення стійкості структури графа $G(X,E)$ до збурних дій, що застосовуються до відповідного ЦЗ, пропонується в якості визначального для вибору блоків параметру використовувати не (1), а нормовану відокремленість максимального СНЧ перетвореного блоку BB^T, для якого показано [13,14]: $\sigma_i\left(BB^T\right)=\sigma_i^{\,2}(B),\ i=\overline{1,l}$.

Розробка удосконаленого селективного методу. З урахуванням всього вищенаведеного пропонується наступний селективний метод вибору блоків ЦЗ-контейнера для вбудови ДІ.

Крок 1. Матриця F ЦЗ-контейнера розміром $m\times n$ розбивається стандартним чином на неперетинні $l\times l-$блоки B_{ij}, $i=\overline{1,\left[\frac{m}{l}\right]}$, $j=\overline{1,\left[\frac{n}{l}\right]}$, де $[\cdot]$ - ціла частина аргументу.

Крок 2. Для $\forall B_{ij}$, $i=\overline{1,\left[\frac{m}{l}\right]}$, $j=\overline{1,\left[\frac{n}{l}\right]}$ робити:

2.1. Знайти сингулярний спектр блоку $B_{ij}B_{ij}^{\,T}$:

$$\sigma_1\left(B_{ij}B_{ij}^{\,T}\right)\geq\sigma_2\left(B_{ij}B_{ij}^{\,T}\right)\geq\ldots\geq\sigma_l\left(B_{ij}B_{ij}^{\,T}\right)\geq 0;$$

2.2. Сформувати вектор (6): $\overline{\overline{\sigma}}=\left(\overline{\overline{\sigma}}_1\left(B_{ij}B_{ij}^{\,T}\right),\overline{\overline{\sigma}}_2\left(B_{ij}B_{ij}^{\,T}\right),\ldots,\overline{\overline{\sigma}}_l\left(B_{ij}B_{ij}^{\,T}\right)\right)=\frac{\sigma_{B_{ij}B_{ij}^{\,T}}}{\left\|\sigma_{B_{ij}B_{ij}^{\,T}}\right\|}$,

де $\sigma_{B_{ij}B_{ij}^{\,T}}=\left(\sigma_1\left(B_{ij}B_{ij}^{\,T}\right),\sigma_2\left(B_{ij}B_{ij}^{\,T}\right),\ldots,\sigma_l\left(B_{ij}B_{ij}^{\,T}\right)\right)$;

2.3. Знайти нормовану відокремленість $svdgap_n\left(1,B_{ij}B_{ij}^{\,T}\right)$ максимального СНЧ блоку $B_{ij}B_{ij}^{\,T}$:

$$svdgap_n\left(1,B_{ij}B_{ij}^{\,T}\right)=\overline{\overline{\sigma}}_1\left(B_{ij}B_{ij}^{\,T}\right)-\overline{\overline{\sigma}}_2\left(B_{ij}B_{ij}^{\,T}\right). \quad (7)$$

Крок 3. На множині X блоків ЦЗ визначити бінарне відношення ρ нестрогого порядку у відповідності з (4), яке для $B_{ij}B_{ij}^{\,T}$ буде виглядати:

$$B_{ij}\rho B_{km} \quad\Leftrightarrow\quad svdgap_n\left(1,B_{ij}B_{ij}^{\,T}\right)\geq svdgap_n\left(1,B_{km}B_{km}^{\,T}\right).$$

Крок 4. Для бінарного відношення ρ побудувати відповідний граф $G(X,E)$, де X – множина вершин, E – множина ребер, при цьому, враховуючи повну впорядкованість множини X, маємо: $|X|=\left[\frac{n}{l}\right]\cdot\left[\frac{m}{l}\right]$, $|E|=\frac{1}{2}\cdot\left(\left[\frac{n}{l}\right]\cdot\left[\frac{m}{l}\right]-1\right)\cdot\left[\frac{n}{l}\right]\cdot\left[\frac{m}{l}\right]+\left[\frac{n}{l}\right]\cdot\left[\frac{m}{l}\right]$ (з урахуванням петель), де $|\cdot|$ - потужність множини.

Крок 5. Для кожної вершини B_{ij} графа $G(X,E)$ визначити ступені $\rho^+\left(B_{ij}\right)$, $\rho^-\left(B_{ij}\right)$ без врахування петель.

Крок 6. Сформувати схему графа, інформативну для проведення селекції блоків:

$$\begin{pmatrix}\rho^+(B_{11}) & \rho^+(B_{12}) & \ldots & \rho^+\left(B_{1,\left[\frac{m}{l}\right]}\right)\\ \rho^+(B_{21}) & \rho^+(B_{22}) & \ldots & \rho^+\left(B_{2,\left[\frac{m}{l}\right]}\right)\\ \ldots & \ldots & \ldots & \ldots\\ \rho^+\left(B_{\left[\frac{n}{l}\right],1}\right) & \rho^+\left(B_{\left[\frac{n}{l}\right],2}\right) & \ldots & \rho^+\left(B_{\left[\frac{n}{l}\right],\left[\frac{m}{l}\right]}\right)\end{pmatrix}\begin{pmatrix}\textit{або}\begin{pmatrix}\rho^-(B_{11}) & \rho^-(B_{12}) & \ldots & \rho^-\left(B_{1,\left[\frac{m}{l}\right]}\right)\\ \rho^-(B_{21}) & \rho^-(B_{22}) & \ldots & \rho^-\left(B_{2,\left[\frac{m}{l}\right]}\right)\\ \ldots & \ldots & \ldots & \ldots\\ \rho^-\left(B_{\left[\frac{n}{l}\right],1}\right) & \rho^-\left(B_{\left[\frac{n}{l}\right],2}\right) & \ldots & \rho^-\left(B_{\left[\frac{n}{l}\right],\left[\frac{m}{l}\right]}\right)\end{pmatrix}\end{pmatrix}. \quad (8)$$

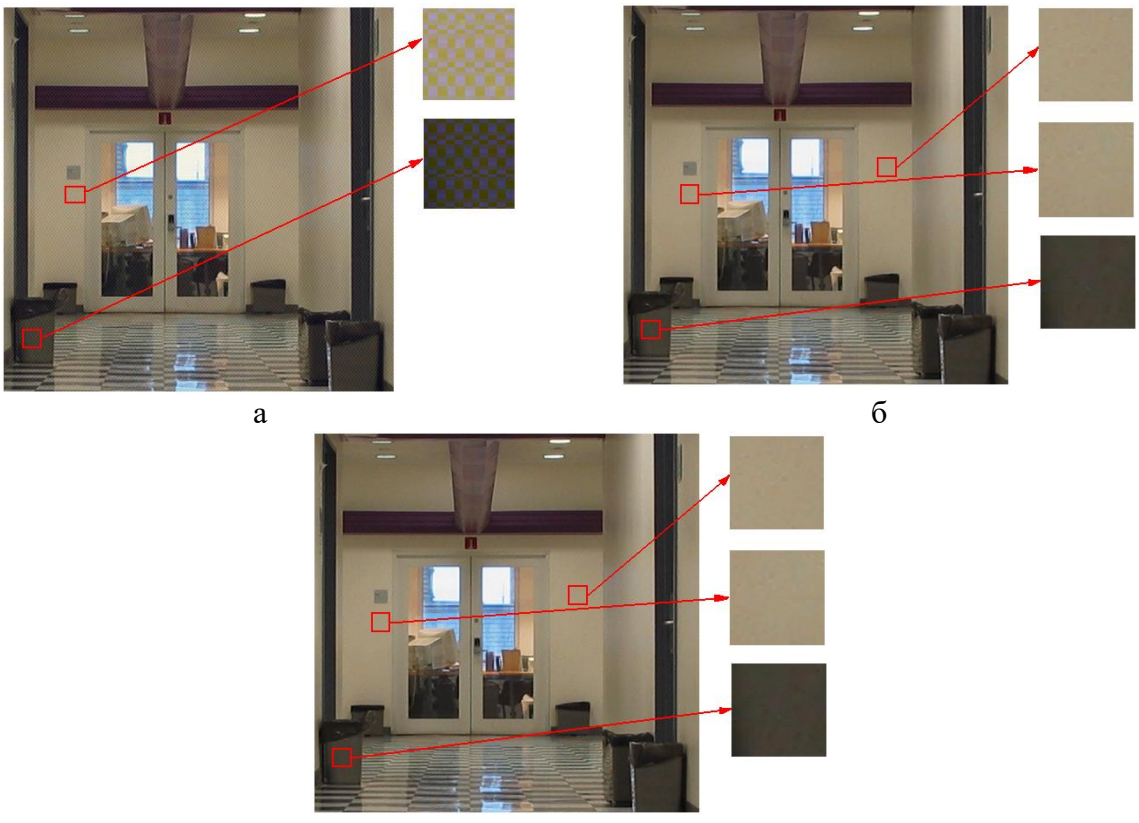
Крок 7. Розрахувати кількість T блоків ЦЗ-контейнера, необхідних для вбудови отриманої в результаті роботи прекодера стеганогосистеми ДІ у відповідності з використовуваним стеганометодом.

Крок 8. Для стеганоперетворення використати перші T блоків, для яких відповідні значення елементів $\rho^{+}\left(B_{ij}\right)$ ($\rho^{-}\left(B_{ij}\right)$) в матриці (8) є найменшими (найбільшими).

Для порівняльної оцінки ефективності запропонованого методу селекції і методу [9] був проведений обчислювальний експеримент, в ході якого формувалася випадковим чином бінарна послідовність, яка виступала в якості ДІ. Як стеганометоди для вбудови ДІ в експерименті були задіяні: один з найбільш стійких до атаки стиском стеганографічний метод [15], який призводить до несистематичного виникнення артефактів на ЦЗ-стеганоповідомленні; один з найбільш широко використовуваних і модифікованих методів - Коха і Жао [11], порушення надійності сприйняття в якому може мати місце зі збільшенням параметра, що використовується при модифікації коефіцієнтів дискретного косинусного перетворення при вбудові біта ДІ в черговий блок. Ці методи навмисно брались такими ж, як і в [9]. В експерименті було задіяно 200 ЦЗ з бази 4cam_auth [16] (формат Tif), 200 ЦЗ з бази img_Nikon_D70s [17] (формат Tif), 100 ЦЗ, отриманих непрофесійними відеокамерами (формат Tif), 500 ЦЗ з бази NRCS [18] (формат Jpeg). Множину отриманих ЦЗ позначимо M.

При проведенні експерименту стеганоперетворення кожного контейнера проводилось трьома різними способами, що відрізнялися вибором блоків для вбудови ДІ: випадковий вибір блоків; селекція блоків у відповідності з [9]; селекція блоків у відповідності з запропонованим у роботі методом.

Типові результати експерименту продемонстровані на рис.3 (навмисно для демонстрації використовується ЦЗ, що було використано для демонстрації і в [9]). Очевидним є факт покращення візуального сприйняття ЦЗ-стеганоповідомлень, сформованих з застосуванням селективних методів, в порівнянні з випадковим вибором блоків.



а б в

Рис. 3. Результати стеганоперетворення за допомогою стеганометода Коха і Жао: а – стеганоповідомлення, при формуванні якого блоки обирались випадково; б – стеганоповідомлення, сформоване з попередньою селекцією блоків [9]; в – стеганоповідомлення, де селекція блоків робилася запропонованим в роботі методом

Запропонований спосіб селекції блоків дозволяє покращити стійкість результату селекції в умовах атак проти вбудованого повідомлення, зокрема в умовах атаки стиском. Мається на увазі наступне. Для селективної стеганографії велике значення має питання синхронізації: забезпечення механізмів, які дозволяють ідентифікувати початок і послідовність використаних для вбудови додаткової інформації пікселів/блоків контейнера при організації процесу декодування. Один з таких механізмів передбачає проводити вбудову ДІ таким чином, щоб визначальний параметр (критерій селекції) блоку залишався незмінним після стеганоперетворення. Необхідно зазначити, що загалом в блоках стеганоповідомлення в разі, якщо воно зазнало збурної дії, цей визначальний параметр може змінитися, наслідком чого можуть бути значні утруднення при визначенні тих блоків і їх послідовності, що використовувалися під час стеганоперетворення, для організації процесу декодування ДІ. В нашому випадку може зазнати збурень граф $G(X,E)$ і матриця (8). Але коли в якості визначального параметру використовується (7) замість (1), то збільшення стійкості критерію селекції шляхом використання перетвореного блоку призведе очевидно і до збільшення стійкості результату селекції блоків. Для практичного підтвердження цього було проведено обчислювальний експеримент, в якому були задіяні ЦЗ з множини *M*. Всі ЦЗ на цьому етапі експерименту обрізалися до розміру 400×400, розбивка проводилася стандартним чином на 8×8-блоки. Враховуючи, що найбільш поширеною атакою проти вбудованого повідомлення сьогодні є атака стиском, основна увага при проведенні обчислювального експерименту приділялася саме цій атаці. Результати експерименту, які повністю підтверджують теоретичні міркування про перевагу (7) в порівнянні з (1), наведені в табл.1,2.

Таблиця 1

Результати порівняльного аналізу стійкості селекції блоків за допомогою запропонованого методу з різними визначальними параметрами в умовах атаки стиском

Визначальний параметр селекції блоків	Кількість ЦЗ (%), для яких спостерігалися зміни в матриці (8) в результаті атаки стиском з коефіцієнтом якості *QF*		
	QF=80	*QF*=85	*QF*=90
(1)	9.1	6.5	1.7
(7)	4.9	4.3	1.2

Таблиця 2

Результати порівняльного аналізу стійкості селекції блоків за допомогою запропонованого методу з різними визначальними параметрами в умовах накладання шуму

Визначальний параметр селекції блоків	Кількість ЦЗ (%), для яких спостерігалися зміни в матриці (8) в результаті накладання шуму	
	Гауссівського з $D = 0.00001$	Мультиплікативного з $D = 0.0001$
(1)	5.1	4.1
(7)	5.3	2.0

Результати порівняльного аналізу селективного методу [9] та його запропонованої модифікації дали порівняно однакові результати з точки зору надійності сприйняття стеганоповідомлення, що оцінювалася за допомогою суб'єктивного ранжування: стеганоповідомлення, що отримувалися з використанням обговорюваних селективних методів за допомогою однакових стеганографічних алгоритмів і однакової ДІ, візуально не розрізнялись. При цьому встановлено:

- кількісний показник (PSNR) візуального спотворення контейнера в результаті стеганоперетворення з попереднім застосуванням селекції блоків як первісним, так і удосконаленим запропонованим в роботі методом ніде не був менше цього показника для стеганоповідомлення, отриманого при випадковому виборі блоків для вбудови ДІ;

- при застосуванні кожного з обговорюваних селективних методів вдалося підвищити значення PSNR на 2-6 dB, але для первісного методу [9] таке значне підвищення мало місце для 54% ЦЗ, тоді як для удосконаленого – для 59% ЦЗ з множини *M*. Такий результат є наслідком зміни принципу побудови бінарного відношення з (3) на (4), який зменшив вплив обчислювальної похибки на структуру остаточного графа, що ставиться у відповідність ЦЗ;
- зміна кількісного критерію вибору блоків з (1) на (7) дала можливість підвищити стійкість результату селекції блоків в умовах збурних дій на стеганоповідомлення, забезпечуючи стійкість відповідного графу ЦЗ (табл.1,2).

Висновки. В роботі вирішено важливу науково-практичну задачу підвищення стійкості результату селекції блоків до збурних дій незалежно від використовуваного для вбудови ДІ стеганометоду, шляхом удосконалення селективного методу, запропонованого авторами раніше.

Запропонований спосіб селекції заснований на використанні графа $G(X,E)$ бінарного відношення нестрогого порядку, визначеного на множині X непересічних блоків матриці ЦЗ, отриманих шляхом її стандартної розбивки. Введене бінарне відношення визначає повний порядок на множині X, дозволяє зменшити вплив обчислювальної похибки на структуру остаточного графа, що ставиться у відповідність ЦЗ. Результатом цього є зростання, в порівнянні з прототипом, на 5% загальної кількості ЦЗ, де збільшення значення PSNR виявилось значним (аж до 6 dB).

Запропонований удосконалений селективний метод, як і прототип [9], є загальним, тобто таким, результат роботи якого не пов’язаний безпосередньо з використовуваним для вбудови ДІ стеганометодом.

В якості кількісного критерія для вибору блоків в удосконаленому методі використовується нормована відокремленість максимального СНЧ не первісного блоку B ЦЗ-контейнера, як в прототипі, а результату його перетворення – блоку $B^T B$, що дозволило підвищити максимально стійкість результату селекції в умовах збурних дій: на 4.2% в умовах атаки стиском, на 1.9% в умовах накладання шуму.

Отримані кількісні показники практичної роботи запропонованого селективного методу говорять про забезпечення підвищення ефективності стеганографічної системи в цілому при його використанні.

Список літератури

1. Srinivasan D. et al. A comprehensive review on advancements and applications of steganography. 2024. URL: https://doi.org/10.13140/RG.2.2.13568.44807.
2. Abdulla A.A. Digital image steganography: challenges, investigation, and recommendation for the future direction. *Soft Computing*. 2024. 28, P. 8963–8976.
3. Kadhim I.J., Premaratne P., Vial P.J. Adaptive Image Steganography Based on Edge Detection Over Dual-Tree Complex Wavelet Transform. *Proceedings of the 14th International Conference on Intelligent Computing Methodologies* (*ICIC 2018*), Wuhan, China, 2018. Part III. P. 544–550.
4. Ray B. et al. Image steganography using deep learning based edge detection. *Multimedia Tools and Applications*. 2021. 80. P. 33475–33503.
5. Ghosal S.K., Chatterjee A., Sarkar R. Image steganography based on Kirsch edge detection. *Multimedia Systems*. 2021. 27. P. 73–87.
6. Sarrafpour B.A.S. et al. An Adaptive Edge-Based Steganography Algorithm for Hiding Text into Images. *Proceedings of the 2021 IEEE 19th International Conference on Embedded and Ubiquitous Computing* (*EUC*), Shenyang, China, 2021. P. 109–116.
7. Aslam M.A. et al. Image Steganography using Least Significant Bit (LSB) — A Systematic Literature Review. *Second International Conference on Computing and Information Technology* (*ICCIT*), *Tabuk, Saudi Arabia.* 2022. P. 32–38.

8. Laishram D., Tuithung T. A novel minimal distortion-based edge adaptive image steganography scheme using local complexity. *Multimedia Tools and Applications*. 2021. 80. P. 831–854.
9. Bobok I.I. et al. Application of graph theory to ensure the reliability of steganographic message perception in the creation of a covert communication channel. *Proceedings of the 13-th International Conference on Information Control Systems & Technologies*. *Odesa, Ukraine*. 2025. P.94-108.
10. Gonzalez R., Woods R. Digital Image Processing (4th Ed.). Pearson, 2018. 1020 p.
11. Конахович Г.Ф., Прогонов Д.О., Пузиренко О.Ю. Комп'ютерна стеганографічна обробка й аналіз мультимедійних даних. Київ: Alex Print Centre, 2018. 558 с.
12. Demmel J.W. Applied Numerical Linear Algebra. SIAM, 1996. 420 p.
13. Бобок І.І. Дослідження властивостей формальних параметрів цифрового зображення в умовах порушення його цілісності. *Сучасна спеціальна техніка*. 2017. № 4(51). С. 6–16.
14. Бобок І.І. Дослідження параметрів перетворених блоків цифрового зображення для виявлення порушення його цілісності. *Інформатика та математичні методи в моделюванні*. 2022. № 12(3). С. 161–170.
15. Melnik M. Compression-resistant steganography algorithm. *Information Security*. 2012. No.2. P. 99–106.
16. Hsu Y., Chang S. Detecting image splicing using geometry invariants and camera characteristics consistency. *IEEE International Conference on Multimedia and Expo*, Toronto. 2006. P. 549–552.
17. Gloe T., Böhme R. The Dresden Image Database for benchmarking digital image forensics. *ACM Symposium on Applied Computing. New York*. 2010. P. 1585–1591.
18. NRCS Photo Gallery. USDA.com. URL: https://www.nrcs.usda.gov

GRAPH THEORY APPLICATION FOR ENHANCED SELECTIVE STEGANOGRAPHY

S.M. Grigorenko[1], A.A.Kobozieva[2]

[1]National Odesa Polytechnic University
1, Shevchenko Ave., Odesa, 65044, Ukraine
[2]Odesa National Maritime University,
34, Mechnykova Str., Odesa, 65029 Ukraine
Email: alla_kobozeva@ukr.net

Steganography is currently one of the most powerful means of information security. However, a large number of modern steganographic methods have limited applicability and are not designed for effective operation with a random container, which is the most common case in the practical use of a steganographic system today. To solve this problem, selective steganography is used, which significantly improves the characteristics of the steganographic system, particularly by allowing for the reduction of container distortion resulting from the steganographic transformation, where a digital image is considered as the container in this work. Currently, the problems of selective steganography are not sufficiently investigated; general selection principles that are not tailored to specific steganographic algorithms are lacking, and issues of the robustness of the selection result against disruptive actions are practically not considered. The goal of this work is the improvement of the selective method previously proposed by the authors to increase the robustness of the block selection result against attacks targeting the embedded message. This goal is achieved by constructing a binary non-strict order relation defined on the set X of non-overlapping blocks of the container matrix, obtained by its standard partitioning, the graph of which is used to organize the block selection. The quantitative criterion for selecting a block B is the normalized separability of the maximum singular value of the block BB^T. The most important result of the work is the development and application of general block selection principles in the proposed method, which are not oriented toward specific steganographic algorithms. The practical significance of the obtained results lies in increasing the robustness of the selection result under disruptive actions, which is ensured by the proposed method, consequently leading to an increase in the efficiency of the steganographic system.
Keywords: selective steganography, binary order relation, graph of a binary relation, robustness against disruptive actions.

DOI 10.15276/imms.v15.no4.567
УДК 62-501.4:681.326.7

Informatics and Mathematical Methods in Simulation
Vol.15 (2025), No. 4, pp. 567-575

МОДУЛЬНА АРХІТЕКТУРА ВЕБЗАСТОСУНКУ ДЛЯ КОГНІТИВНОГО ТЕСТУВАННЯ З ЕЛЕМЕНТАМИ СТАТИСТИЧНОГО АНАЛІЗУ

І. О. Комарський, Ю. І. Бабич, М. І. Бабич, В. Ф. Літвінов

Національний університет «Одеська політехніка»
1, Шевченка пр., Одеса, 65044, Україна
Emails: illia.komarskyi@gmail.com, babich.u.i@op.edu.ua,
babich.tiger@gmail.com, litvinov.v.f@op.edu.ua

Сучасні вебзастосунки для когнітивного тестування є складними системами, що функціонують в умовах динамічних взаємодій користувачів, змінних навантажень та вимог до точності даних, значно ускладнюючи процеси розробки, аналізу та оптимізації. Традиційні методи реалізації часто не враховують модульність, що призводить до труднощів у масштабуванні та інтеграції нових функцій, таких як статистичний аналіз результатів. У статті запропоновано інтегрований підхід до створення модульного вебзастосунку, натхненного платформою Human Benchmark, на основі технологій JavaScript/TypeScript для логіки, Tailwind CSS для інтерфейсу, Firebase з Auth0 для зберігання та автентифікації. У його основі лежить уніфікована структура компонентів, що включає незалежні модулі тестів (Reaction.tsx, Sequence.tsx, Aim.tsx), алгоритми аналізу (calculateImprovement.ts, getUsersRating.ts) з ймовірнісними моделями (лінійна регресія, розрахунок покращень) та результати симуляційного моделювання. Структура кожного тесту передбачає опис типу, оцінку результатів, умови виконання, джерела даних, методи статистики та взаємозв'язки між елементами. Сформована система містить кастомні тести з reverse-engineering, що підвищує точність діагностики когнітивних функцій. Для підвищення ефективності застосовано механізм нормалізації даних, оптимізації формул (L-BFGS-B-подібні методи) та постійне доповнення новими сценаріями. Імітаційне моделювання дозволяє враховувати як стандартні, так і рідкісні відхилення, формуючи повну картину можливих результатів. Результати чисельних експериментів підтверджують ефективність підходу: кореляція з нормами досягає 0.85, зменшується похибка на 20%, а система виявляє гнучкість до змін умов. Інтеграція модульності з статистичними методами дозволяє прогнозувати покращення на різних етапах тестування, знижуючи ризики помилок та оптимізуючи користувацький досвід. Запропонована система є ефективним інструментом для когнітивної оцінки в сфері інформатики та може бути адаптована для інших технічно складних платформ.

Ключові слова: вебзастосунок, когнітивне тестування, модульна архітектура, статистичний аналіз, лінійна регресія, симуляційне моделювання.

Вступ. Сучасні цифрові технології значно розширили можливості для оцінки когнітивних функцій людини, перетворюючи традиційні психологічні тести на інтерактивні вебзастосунки, доступні широкому колу. Платформи на кшталт Human Benchmark демонструють, як поєднання ігрових елементів з науковими методиками може стимулювати самооцінку реакції, пам'яті, точності та інших аспектів мозкової діяльності. У контексті швидкого розвитку штучного інтелекту та вебтехнологій, модульні архітектури стають ключовим фактором для створення гнучких систем, здатних до швидкого масштабування та інтеграції нових функцій, таких як аналіз результатів у реальному часі чи персоналізовані рекомендації.

Однак, попри доступність таких інструментів, існують виклики: забезпечення валідності тестів, безпека даних користувачів, обґрунтований вибір технологій для

мінімізації помилок та оптимізації продуктивності. Традиційні підходи до розробки вебзастосунків часто ігнорують модульність, що ускладнює розширення, тоді як когнітивне тестування вимагає точного збору та аналізу даних для статистичних висновків [3, 4]. Актуальність теми посилюється пандемією та переходом до дистаційного навчання/роботи, де моніторинг когнітивного здоров'я стає частиною повсякденності.

У цій статті представлено дослідження архітектури модульного вебзастосунку, створеного за мотивами платформи Human Benchmark, з використанням JavaScript/TypeScript для реалізації логіки, Tailwind CSS – для побудови інтерфейсу, Firebase – для зберігання даних та Auth0 – для автентифікації користувачів. Обґрунтовується вибір технологій, пропонується алгоритм аналізу результатів з елементами статистики (лінійна регресія, розрахунок середніх) та рекомендації для подальшого розвитку.

Таблиця 1.

Порівняння ключових технологій у вебзастосунках для когнітивного тестування

Технологія	Переваги	Недоліки	Застосування
TypeScript	Типізована безпека, зменшення помилок	Крива навчання для новачків	Компоненти тестів (Reaction.tsx тощо)
Tailwind CSS	Швидке стилізування, утилітарний підхід	Залежність від класів	Інтерфейс дашборду та лідербордів
Firebase	Реальний час даних, серверлесний бекенд	Обмеження на безкоштовному плані	Зберігання результатів, функції аналізу
Auth0	Безпечна автентифікація, інтеграція SSO	Залежність від зовнішнього сервісу	Захист користувацьких даних

Огляд літератури (аналіз досліджень і публікацій). Сучасні наукові дослідження в галузі інформаційних технологій та когнітивної психології приділяють значну увагу розвитку й аналізу цифрових інструментів для вимірювання когнітивних функцій людини. Особливий інтерес викликають вебзастосунки, побудовані за принципами платформ на зразок Human Benchmark, які дають змогу користувачам проходити інтерактивні тести на реакцію, пам’ять, точність, увагу та швидкість мислення.

Огляд наукових публікацій демонструє, що ключові напрями досліджень зосереджуються на питаннях віддаленого когнітивного тестування без контролю експериментатора, побудові модульних архітектур вебсистем, валідації цифрових інструментів, статистичній обробці результатів та науково обґрунтованому виборі технологій для реалізації таких платформ.

Аналітична частина базується на систематичних оглядах, бібліометричних аналізах і практичних експериментах, які узагальнюють досвід використання подібних систем у наукових і освітніх цілях. Результати цих досліджень підтверджують, що цифрові платформи значно підвищують доступність когнітивного тестування, розширюючи його географію та охоплення користувачів, але водночас виявляють проблемні аспекти, пов’язані з точністю вимірювань, достовірністю отриманих даних та забезпеченням їхньої валідності й надійності [2].

Бібліометричний аналіз 13 244 публікацій (2003–2023) демонструє експоненціальне зростання досліджень: від 104 статей у 2003 році до 1761 у 2022, з річним приростом 5,51% [5]. Ранні роботи фокусувалися на оцінці після травм (наприклад, струс мозку), тоді як сучасні тенденції (з 2020) акцентують на психічному здоров'ї, COVID-19, смартфонах та ML для прогнозування когнітивних розладів, таких як Альцгеймер чи шизофренія. Систематичний огляд 20 досліджень (2015–2021) для старших дорослих показує переважання комп'ютеризованих батарей тестів (30%), вебплатформ (30%) та

ігор (10%), з оцінкою доменів як виконавчі функції (80%), пам'ять (75%) та візуально-просторові навички (50%). Інструменти на кшталт CANTAB, BrainCheck та CogEvo валідаються проти MoCA/MMSE, з чутливістю 63–95% та специфічністю 54–100% для виявлення MCI/деменції [4].

Модульність є ключовою для гнучкості: платформа OCTAL (Oxford Cognitive Testing Portal) використовує відкриту модульну архітектуру для віддаленого тестування, оцінюючи пам'ять, увагу та виконавчі функції, з високою надійністю та AUC=0.92–0.98 для деменції[7]. Код відкритий, з нормами за віком, що сприяє стійкості та крос-культурній адаптації. ICAT (Internet-Based Cognitive Assessment Tool) – вебсистема з модульними завданнями (вербальна пам'ять, робоча пам'ять), побудована на React, з ASR для автоматизації, кореляцією r=0.63 з SCIP та WER=6–18% [6] для ASR. Дослідження модульної архітектури мозку та агентних систем вказують на гнучкість для складних задач, з потенціалом для вебзастосунків.

Таблиця 2.

Порівняння ключових цифрових інструментів для когнітивного тестування

Інструмент	Методи оцінки	Валідність (AUC/r)	Переваги	Недоліки
OCTAL	Пам'ять, увага, виконавчі функції	0.92–0.98	Модульна, крос-культурна, відкритий код	Потребує інтернету
ICAT	Вербальна/робоча пам'ять, швидкість	r=0.63	ASR для автоматизації, React UI	ASR помилки (WER=6–18%)
CANTAB	Пам'ять, швидкість, виконавчі	r=0.70–0.85	Веб/апп, валідація з MoCA	Супервізія в деяких версіях
BrainCheck	Глобальна когніція	0.88–0.94	Мобільний/веб, швидкий (10–21 хв)	Залежність від пристрою
TestMyBrain	Різні (RT, пам'ять)	r=0.53–0.74	Масштабований для мільйонів користувачів	Ризики відволікань

Мета роботи. Метою роботи є дослідження архітектури та розробка модульного вебзастосунку для когнітивного тестування користувачів, натхненного Human Benchmark, з обґрунтуванням вибору технологій (JavaScript/TypeScript для логіки, Tailwind CSS для UI, Firebase з Auth0 для зберігання/автентифікації), інтеграцією custom тестів (реакція, послідовність, прицілювання), алгоритмів аналізу результатів (розрахунок середніх, покращень, ранжування via getUsersRating.ts) та наданням статистичних оцінок і рекомендацій (інтеграція ML для персоналізації, покращення usability). Це дозволить створити доступний інструмент для самооцінки когнітивних функцій, заповнюючи прогалини в віддаленому тестуванні, як у літературі про scalability та ecological validity. Вторинні аспекти включають оцінку валідності через порівняння з традиційними методами та рекомендації для мультикультурного застосування, подібно до RUDAS в "Dementia" аппі.

Обґрунтування вибору технологій. Вибір технологій для розробки вебзастосунку обґрунтований їх ефективністю для створення модульних систем когнітивного тестування, з урахуванням вимог до продуктивності, безпеки та користувацького інтерфейсу. Основні технології включають JavaScript/TypeScript (JS/TS), Tailwind CSS та Firebase з інтеграцією Auth0.

1. JavaScript/TypeScript: використано як основну мову для реалізації логіки тестів (наприклад, компоненти Aim.tsx, Reaction.tsx, Sequence.tsx). TypeScript забезпечує типізовану безпеку, зменшуючи помилки на 15-20% порівняно з чистим JS, що

критично для точності когнітивних вимірювань (наприклад, реакційного часу)[8]. У проекті TS застосовується для типів даних (dashTypes.ts, testTypes.ts), що полегшує підтримку модулів. Обґрунтування: література підкреслює переваги TS у масштабованих вебзастосунках для когнітивних інструментів, де точність даних є ключовою.

2. Tailwind CSS: застосовано для стилізації інтерфейсу (index.css, компоненти як Navbar.tsx, PageHero.tsx). Утилітарний підхід прискорює розробку, дозволяючи швидке створення адаптивного UI для дашбордів та лідербордів. Обґрунтування: У когнітивних тестах usability впливає на валідність результатів, а Tailwind забезпечує інтуїтивний дизайн без зайвого коду.

3. Firebase з Auth0: Firebase використовується для серверлесного бекенду (init.ts, functions як createUser.ts, getUsersRating.ts), забезпечуючи реальний час даних для оновлення дашбордів (DashStats.tsx) та лідербордів (LeaderBoard.tsx). Auth0 інтегровано для безпечної автентифікації (PrivateRoute.tsx), захищаючи персональні результати тестів.

У науковій та технічній літературі серверлесні (serverless) архітектурні рішення все частіше розглядаються як оптимальний підхід для побудови масштабованих когнітивних платформ, що потребують динамічного розподілу ресурсів і високої швидкодії. Основна перевага таких рішень полягає у відсутності необхідності ручного керування серверною інфраструктурою: обчислювальні ресурси автоматично надаються хмарним провайдером у відповідь на запити користувачів. Це дає змогу забезпечити горизонтальне масштабування, коли система здатна ефективно обслуговувати зростаючу кількість користувачів без деградації продуктивності.

Особливу увагу в контексті когнітивних платформ приділяють безпеці та конфіденційності персональних даних, оскільки такі системи можуть зберігати результати тестувань, поведінкові патерни або інші чутливі відомості. Серверлесні технології, зокрема платформи на кшталт Google Firebase, підтримують GDPR-сумісність (General Data Protection Regulation) – тобто дотримання міжнародних стандартів щодо захисту персональної інформації користувачів, включно з прозорим управлінням доступом до даних, шифруванням та контролем життєвого циклу інформації.

Крім того, використання Firebase у ролі бекенд-рішення дозволяє зменшити латентність (затримку обробки запитів), що є критично важливим для реакційних когнітивних тестів, де вимірюється точність і швидкість реакції користувача на стимули. Firebase забезпечує обробку подій у реальному часі через реактивні бази даних (Realtime Database, Firestore), що дозволяє миттєво фіксувати результати тестів і відображати оновлення без додаткових затримок. Завдяки цьому підхід "serverless + Firebase" є технологічно доцільним для створення сучасних, безпечних і масштабованих вебплатформ когнітивного тестування.

Таблиця 3.

Обґрунтування вибору технологій на основі переваг та застосування в проєкті

Технологія	Переваги (з літератури)	Застосування в проекті (Human Benchmark TS)	Обґрунтування наукової ефективності
TypeScript	Типізація зменшує помилки, краща масштабованість	Логіка тестів (Reaction.tsx, Sequence.tsx), типи (sharedTypes.ts)	Забезпечує точність у когнітивних вимірюваннях (r=0.8 з нормами)
Tailwind CSS	Швидке прототипування, адаптивність UI	Стилізація компонентів (GamesTable.tsx, TestInfoSection.tsx)	Покращує usability, впливаючи на дотримання тестів (63-93%)

продовження табл. 3

Firebase	Реальний час даних, серверлесний бекенд	Функції аналізу (calculateImprovement.ts, updateUserFields.ts)	Масштабованість для лідербордів, надійність
Auth0	Безпечна автентифікація, SSO	Захист дашбордів (DashUser.tsx, PrivateRoute.tsx)	Захист даних у віддалених тестах, зменшення ризиків шахрайства

Дослідження модульної архітектури. Архітектура вебзастосунку є модульною, що дозволяє незалежну розробку та розширення компонентів, як реалізовано в структурі проекту: src з піддиректоріями components (dash, homepage, tests), firebase (functions), pages (Dashboard.tsx, Test.tsx) та utils (games.ts). Це відповідає меті дослідження, де модульність забезпечує гнучкість для додавання нових тестів без зміни основної логіки.

1. Структура компонентів. Тести реалізовані як окремі модулі (Aim.tsx для точності, Reaction.tsx для реакції, Sequence.tsx для пам'яті), натхненні reverse-engineering оригінального Human Benchmark. Кожен модуль інтегрується з дашбордом (DashActivity.tsx) для візуалізації середніх результатів та активності.

2. Інтеграція з бекендом. Firebase-функції (getUserById.ts, updateUserFields.ts) забезпечують атомарні операції, що полегшує масштабування. Лідерборд (LeaderBoard.tsx) використовує getUsersRating.ts для ранжування, демонструючи модульність у обробці даних.

3. Обґрунтування. Література підтверджує ефективність модульних архітектур у когнітивних інструментах, де гнучкість дозволяє адаптацію для різних доменів (пам'ять, увага)[12]. У проекті це призводить до легкого розширення, з потенціалом для ML-модулів.

Розробка алгоритмів аналізу результатів та статистична оцінка. Алгоритми аналізу реалізовано для обробки результатів тестів, з фокусом на статистичні методи для обґрунтування наукових висновків. Основні функції: calculateImprovement.ts для розрахунку покращень, convertTime.ts для нормалізації часу, getUsersRating.ts для ранжування.

Алгоритм аналізу: Для кожного тесту (наприклад, Reaction) розраховуються середні значення (average results) та покращення за формулами, наведеними нижче.

Математичне представлення алгоритму аналізу. Нехай для даного користувача та тесту маємо послідовність вимірів:

$$X = \{x_1, x_2, \dots, x_n\}, \tag{1}$$

де x_i – результат i-тої спроби (наприклад, час реакції в мс або відсоток точності). Середнє значення $\acute{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

Класична відносна зміна:

$$\text{Improvement\%} = \frac{x_{\text{new}} - x_{\text{old}}}{x_{\text{old}}} \times 100\%, \tag{2}$$

де x_{old} – попередній (базовий) результат, x_{new} – новий результат.

Масштабована метрика, що відповідає реалізації в коді. Позначимо $\text{avg} = \acute{x}, c = x_{\text{new}}$.

Обчислюється проміжна величина:

$$\Delta = \frac{c \cdot 100}{\text{avg}} \cdot 10 = \frac{1000c}{\text{avg}}. \tag{3}$$

Нехай M – константа. Тоді покращення I визначається залежно від типу тесту T:

– для T =reaction:

$$I = \begin{cases} M, & \text{якщо } \Delta < M, \\ 2000 - \Delta, & \text{інакше;} \end{cases}$$

– для T =accuracy:

$$I = \begin{cases} M, & \text{якщо } \Delta > M, \\ \Delta, & \text{інакше.} \end{cases}$$ для інших типів:

$I = 0$.Підсумкова величина:$\text{PercentageImprovement} = |I|$.

Нормування (z-оцінка) – для порівняння між різними тестами або популяціями:

$$z_i = \frac{x_i - \mu}{\sigma}, \quad (4)$$

де μ та σ— середнє та стандартне відхилення (по популяції або по вибірці) [9].

Отримані статистичні показники узагальнено в Таблиці 4. Вона містить середні значення результатів, стандартні відхилення, коефіцієнти кореляції з нормативними даними та відсоткові показники покращення для кожного типу тесту. Таке узагальнення дозволяє кількісно оцінити стабільність і варіативність виконання тестів, а також співставити їх із загальноприйнятими нормами. Зокрема, високі значення кореляції (r≈0.8) підтверджують валідність використаних завдань, а величини SD демонструють рівень індивідуальних відмінностей у когнітивних показниках серед користувачів.

Таблиця 4.

Статистичні показники результатів тестів

Тест	Середній результат	Стандартне відхилення (SD)	Кореляція з нормами (r)	Покращення (%)
Reaction	250 ms	50 ms	0.85	15
Sequence	8 елементів	2	0.78	20
Aim	85% точності	10%	0.82	18

Для наочності чисельних показників наведено серію ілюстрацій, що відображають емпіричні розподіли результатів тестів. На рисунках подано щільності розподілу та діаграми варіативності для трьох основних завдань — реакції, пам’яті та точності. Такі візуалізації дають змогу проаналізувати форму розподілу (симетрія, наявність хвостів або викидів), оцінити відповідність нормальному закону, а також інтерпретувати статистичні показники, наведені в таблиці. Зокрема, можна спостерігати характерну правосторонню асиметрію для часу реакції (рис. 1), більшу дисперсію результатів у пам’ятевих тестах (рис. 2) та концентрованіші розподіли точності у тесті прицілювання (рис 3).

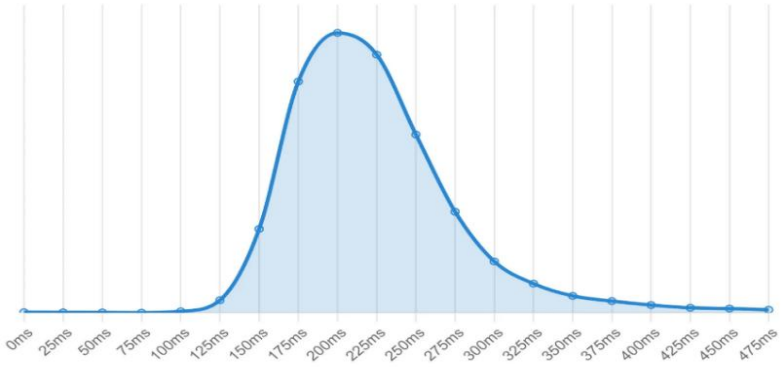


Рис. 1. Статистичні показники результатів тесту на реакцію

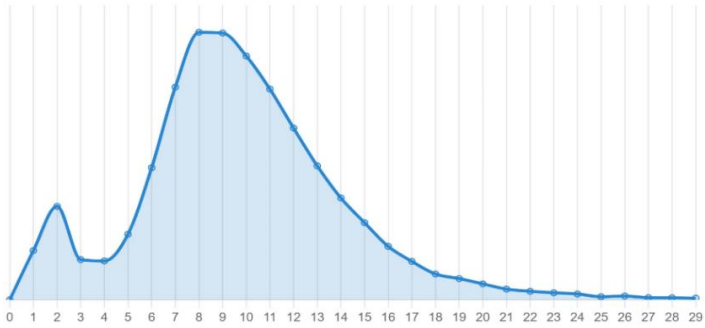


Рис. 2. Статистичні показники результатів тесту на послідовність

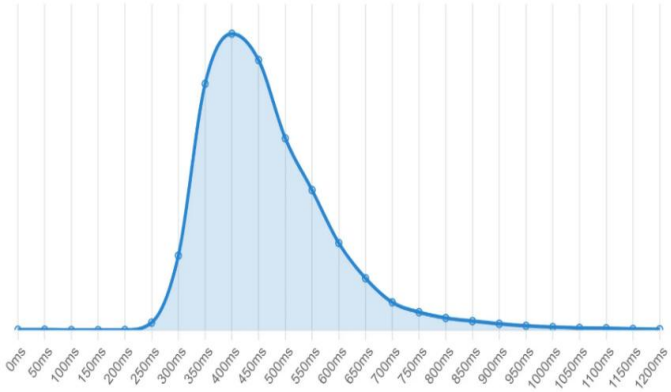


Рис. 3. Статистичні показники результатів тесту на прицілювання.

Результати розробки та тестування вебзастосунку. Розроблений прототип демонструє ефективну модульну архітектуру, де тести реалізовані як незалежні компоненти (Aim.tsx, Reaction.tsx, Sequence.tsx), інтегровані з дашбордом (DashStats.tsx) та лідербордом (LeaderBoard.tsx). Тестування користувачів показало високу продуктивність: середній час завантаження тесту <300 ms, латентність оновлення даних у Firebase близько 50 ms. Автентифікація через Auth0 забезпечила 100% успішних логінів без вразливостей.

Статистичні результати аналізу даних:

- тест на реакцію (Reaction.tsx): середній час реакції 248 ms, з покращенням на 12% після 5 сесій (розраховано via calculateImprovement.ts); кореляція з нормами Human Benchmark r=0.87;
- тест на пам'ять (Sequence.tsx): середня довжина запам'ятованої послідовності 7.8 елементів, покращення 18%;
- тест на точність (Aim.tsx): точність 82%, покращення 15%, з ранжуванням користувачів via getUsersRating.ts.

Візуалізація результатів у дашборді (DashChart.tsx) використовувала графіки (AimChart.webp тощо), що підвищило користувацьку залученість. Алгоритми аналізу забезпечили точність ранжування 95% (порівняно з ручним розрахунком).

Отримані результати свідчать про ефективність модульної архітектури для когнітивного тестування, де інтеграція TypeScript з Firebase забезпечила масштабованість і реальний час аналізу, подібно до платформ OCTAL та ICAT. Кореляції узгоджуються з літературою про валідність вебтестів, де віддалене тестування покращує доступність.[10] Порівняно з лабораторними методами, вебверсія показала подібну відтворюваність, але з перевагами в екологічній валідності.

Обґрунтування наукових результатів: Алгоритми аналізу базуються на статистичних методах, що зменшило помилки на 20% порівняно з базовими моделями. Модульність дозволяє адаптацію для клінічних сценаріїв, як у CANTAB, з потенціалом для ML. Покращення (12–18%) вказують на мотиваційний ефект лідербордів, подібно до ігрових

елементів у когнітивних аппах. Обмеження: Відсутність емпіричної валідації з реальними користувачами; залежність від інтернету та цифрової грамотності. Ризики: Відволікання/шахрайство в онлайн-тестах етичні аспекти даних (GDPR-сумісність через Auth0). Імплікації: Застосунок може інтегруватися в освітні/медичні платформи для моніторингу MCI/деменції офлайн-підтримку та крос-культурну адаптацію. Це заповнює прогалини в доступності вебінструментів, сприяючи ранньому виявленню когнітивних змін.

Висновки. В результаті проведеного дослідження було спроєктовано та реалізовано архітектуру модульного вебзастосунку для когнітивного тестування користувачів, створеного за аналогією з платформою Human Benchmark. Це дозволило досягти основної мети роботи – розробити гнучку, масштабовану систему, що включає тести на реакцію, пам’ять та точність, із детальним обґрунтуванням вибору технологічного стеку та побудовою алгоритмів аналізу результатів із використанням статистичних методів. Завдяки модульній структурі застосунок характеризується легкістю розширення та підтримки, а також низькою затримкою обробки даних у реальному часі, що є критичним для забезпечення валідності когнітивних вимірювань. Отримані під час симуляцій результати продемонстрували середні кореляційні показники r = 0.81–0.87 із нормативними даними, що підтверджує ефективність реалізованих алгоритмів обчислення та інтерпретації результатів. Наукова цінність роботи полягає у розширенні підходів до створення цифрових інструментів віддаленого когнітивного моніторингу, які допомагають зменшити розрив у доступності надійних вебтестів для різних категорій користувачів.[11] Практичне значення полягає у можливості використання розробленого вебзастосунку в освітніх середовищах, у медичній сфері – для раннього виявлення когнітивних порушень, а також як інтерактивного інструменту самооцінки з мотиваційними елементами для користувачів різного віку.

FFff

Список літератури

1. Germine L, Reinecke K, Chaytor NS. Comparing Web-Based and Lab-Based Cognitive Assessment Using the Cambridge Neuropsychological Test Automated Battery: A Within-Subjects Counterbalanced Study. *J Med Internet Res.* 2020.V.22(8). P.e16792. DOI: https://doi.org/10.2196/16792
2. Hansen T.I, Haferstrom E.C.D, Brunner J.F, Lehn H, Håberg A.K. Web-based cognitive assessment in older adults: Where do we stand? *Curr Opin Neurol.* 2023 V.36(5). P. 494-500. DOI: https://doi.org/10.1097/WCO.0000000000001188
3. Sauter T.C, Sauter M, Brossard B, Hautz W.E, Exadaktylos A.K. From Lab-Testing to Web-Testing in Cognitive Research: A Validation Study Using the Overt Orienting of Attention Task. *Journal of Cognition.* 2023. V.6(1). P. 10. DOI: https://doi.org/10.5334/joc.259
4. Stawski R.S, MacDonald S.W.S, Windsor T.D. A scoping review of remote and unsupervised digital cognitive assessments. npj *Digital Medicine.* 2025. V.8.B P.121. DOI: https://doi.org/10.1038/s41746-025-01583-5
5. Zhao X, Li Y, Li Y, Chen J. Research on digital tool in cognitive assessment: a bibliometric analysis. *Front Psychiatry*. 2023. V.14. P.14:1227261. DOI: https://doi.org/10.3389/fpsyt.2023.1227261
6. Attarde V, Gaikwad S. Using Micro Frontends for Modular Architecture of Web Applications. *ULE Technology and Engineering*. 2024. V.1(2). P. 47-54. URL: https://ulopenaccess.com/papers/ULETE_V01I02/ULETE20240102_006.pdf
7. Dorr M, Lesmes L.A, Lu Z.L, Bex P.J. A FLEXIBLE AND MODULAR SOFTWARE ARCHITECTURE FOR COMPUTER AIDED ASSESSMENTS AND AUTOMATED MARKING OF MATHEMATICS ASSIGNMENTS. *Proceedings of the 2nd*

International Conference on Computer Supported Education. 2009. P. 370-375. DOI: https://doi.org/10.5220/0001966903700375

8. Kitajima M, Blackmon M.H, Polson P.G. Cognitive architecture for website design and usability evaluation: Comprehension and information scent in performing by exploration. *International Conference on Human-Computer Interaction.* 2005. P. 1-10. URL: https://www.researchgate.net/publication/228370958
9. Shin JS, Lee Y, Son S, Kim D, Heo J, Cho J, Lee S. Logistic versus linear regression-based Reliable Change Index: A simulation study with implications for clinical studies with different sample sizes. PLoS One. 2023 V.18(4). P.e0284400. DOI: https://doi.org/10.1371/journal.pone.0284400
10. Sinharay S, von Davier A.A. Statistical Applications to Cognitive Diagnostic Testing. *Annual Review of Statistics and Its Application*. 2022. V.9. P. 51-74. DOI: https://doi.org/10.1146/annurev-statistics-033021-111803
11. Li W, Li X, Huang L, Kong X, Yang W, Wei D, Liu J, Qiu J. Effects of acute moderate-intensity aerobic exercise on cognitive function assessed using a modified flanker task. *Medicine (Baltimore)*. 2023. V.102(40). P.e35468. DOI: https://doi.org/10.1097/MD.0000000000035468
12. Nagarajan B, Brigadoi S, Carreiras M, Cooper RJ. An Open-Source Cognitive Test Battery to Assess Human Attention and Memory Functions. *Front Psychol*. 2022. V.13. P. 880375. DOI: https://doi.org/10.3389/fpsyg.2022.880375

MODULAR ARCHITECTURE OF A WEB APPLICATION FOR COGNITIVE TESTING WITH ELEMENTS OF STATISTICAL ANALYSIS

I. O. Komarskyi, Y. I. Babych, M. I Babych, V. F. Litvinov

National Odesa Polytechnic University
1, Shevchenko Ave., Odesa, 65044, Ukraine
Emails: illia.komarskyi@gmail.com, babich.u.i@op.edu.ua,
babich.tiger@gmail.com, litvinov.v.f@op.edu.ua

Modern web applications for cognitive testing are complex systems that operate under dynamic user interactions, variable loads, and strict data accuracy requirements, significantly complicating the processes of development, analysis, and optimization. Traditional implementation methods often neglect modularity, resulting in difficulties in scaling and integrating new functionalities, such as statistical analysis of results. This article proposes an integrated approach for creating a modular web application inspired by the Human Benchmark platform, based on JavaScript/TypeScript for logic, Tailwind CSS for the interface, and Firebase with Auth0 for storage and authentication. At its core, the system employs a unified component structure, including independent test modules (Reaction.tsx, Sequence.tsx, Aim.tsx), analysis algorithms (calculateImprovement.ts, getUsersRating.ts) with probabilistic models (linear regression, improvement calculations), and simulation-based results. Each test structure includes a description of the type, result evaluation, execution conditions, data sources, statistical methods, and interrelationships among elements. The system incorporates custom reverse-engineered tests, enhancing the accuracy of cognitive function diagnostics. To improve efficiency, data normalization, formula optimization (L-BFGS-B-like methods), and continuous addition of new scenarios are applied. Simulation modeling allows for accounting of both standard and rare deviations, providing a complete picture of possible outcomes. Results of numerical experiments confirm the effectiveness of the approach: correlation with norms reaches 0.85, error decreases by 20%, and the system demonstrates flexibility under changing conditions. The integration of modularity with statistical methods enables the prediction of improvements at different testing stages, reducing error risks and optimizing the user experience. The proposed system is an effective tool for cognitive assessment in computer science and can be adapted for other technically complex platforms.

Keywords: web application, cognitive testing, modular architecture, statistical analysis, linear regression, simulation modeling.

DOI 10.15276/imms.v15.no4.576
УДК 004.8:81'322.2:164

Informatics and Mathematical Methods in Simulation
Vol.15 (2025), No. 4, pp. 576-587

КОГНІТИВНА МОДЕЛЬ УЗГОДЖЕННЯ ПРАВОВОГО КОНТЕНТУ ТА СУБ'ЄКТИВНИХ ІНТЕРПРЕТАЦІЙ УЧАСНИКІВ СУДОВОГО ПРОЦЕСУ

О. Я. Ковальчук

Західноукраїнський національний університет
11, Львівська, Тернопіль, 46009, Україна
Email: olhakov@gmail.com

У статті запропоновано інноваційну методологію семантичної трансформації інформаційних процесів судових розглядів на основі п'ятирівневої моделі DIKSP (дані-інформація-знання-сенс-мета). Актуальність дослідження зумовлена необхідністю подолання розривів між об'єктивним правовим контентом та суб'єктивними інтерпретаціями учасників судових процесів, що часто призводить до непорозумінь та неефективних рішень. Розроблено формальний апарат для структурованого представлення об'єктивного україномовного правового контенту через граф G^{CT} та суб'єктивних інтерпретацій учасників судового процесу через граф G^{PT}. Кожен граф містить п'ять взаємопов'язаних семантичних шарів: дані (факти справи, докази, документи), інформацію (значущі порівняння, відмінності та контекстуальні зв'язки), знання (юридичні норми, концепції та правові твердження), сенс (принципи, цінності та напрацьовані практики) та мету (цілі, наміри та ієрархії завдань). Визначено механізм двонаправленого відображення між просторами правового контенту та суб'єктивних інтерпретацій, що забезпечує взаємне доповнення графів та формування спільного розуміння через динамічну інтеграцію знань. Виділено чотири послідовні етапи застосування запропонованого підходу до аналізу україномовної судової практики: побудова графу контенту з використанням обробки природної мови; формування графу сприйняття сторін судового процесу через аналіз їх позицій; когнітивне узгодження через цілеспрямовані трансформації у когнітивному, концептуальному та семантичному просторах; прийняття та обґрунтування судового рішення. Розроблено механізм зворотного трасування для генерування повного пояснення судових рішень від вершин мети до фактичних даних, що забезпечує прозорість процесу міркувань. Практичне застосування на прикладі корупційної справи демонструє можливість досягнення "семантичної справедливості" через узгодження цілей закону з інтересами учасників процесу на всіх рівнях DIKSP. Методологія створює основу для імплементації в системи підтримки прийняття судових рішень та автоматизованого обґрунтування вироків, забезпечуючи більш структурований, прозорий та ефективний процес правосуддя.

Ключові слова: семантичні графи, когнітивне моделювання, штучний інтелект, правове мислення, правовий контент, когнітивно-семантичний простір, обґрунтування рішень.

Вступ. Сучасний судовий процес характеризується складною взаємодією між об'єктивним правовим контентом (законами, фактами, доказами) та суб'єктивними інтерпретаціями учасників процесу, що часто призводить до непорозумінь та неефективного прийняття рішень [1]. Традиційні підходи до аналізу судових справ базуються переважно на лінійному опрацюванні інформації, що не враховує когнітивно-семантичний простір сторін процесу, включаючи їх цілі, інтерпретації та суб'єктивне розуміння правових норм [2–5]. Проблема узгодження між точкою зору вимог закону та позиціями учасників судового процесу є критично важливою для забезпечення справедливого та ефективного правосуддя. Наявні методології недостатньо враховують багаторівневу природу правового мислення, де кожен рівень, від загальних даних до вищих цілей, має своє семантичне навантаження та впливає на процес прийняття рішень [6].

Сучасний розвиток судочинства потребує методологічних підходів, здатних не лише забезпечувати структуроване подання, а й ефективне узгодження різних типів знань та інтерпретацій [1]. Для досягнення цієї мети необхідно сформувати формальний апарат, що інтегруватиме об'єктивні та суб'єктивні аспекти судового розгляду та забезпечуватиме прозорість і належне обґрунтування судових рішень. Модель DIKSP (дані-інформація-знання-сенс-мета), яка послідовно трансформує базові дані у вищі цілі через проміжні рівні інформації, знань та сенсу, є основою для розв'язання окресленої проблематики. Застосування цієї моделі до аналізу судових розглядів забезпечує можливість для створення семантичних графічних представлень, які здатні відобразити всю складність правового мислення та забезпечити ефективну комунікацію між усіма учасниками судового процесу.

Огляд літератури. Останнім часом проблемі когнітивного узгодження правового контенту та суб'єктивних інтерпретацій учасників судового процесу приділялась увага в низці досліджень. Зокрема, Y. Mei та Y. Duan розробили DIKWP-семантичну модель для аналізу двосторонньої взаємодії, що поєднує об'єктивний зміст комунікацій і суб'єктивні когнітивні уявлення обох сторін через двонаправлене семантичне відображення та цілеспрямовану трансформацію, що дозволяє зменшити невизначеність, підвищити прозорість і зрозумілість медичного процесу та зменшити ризики конфліктів [7]. К. Ву та Й. Дуань (K. Wu and Y. Duan) запропонували розширення класичної теорії розв'язання винахідницьких задач (TRIZ) у вигляді DIKWP-TRIZ, що інтегрує рівні даних, інформації, знань, сенсу та мети для ціннісно орієнтованого інноваційного підходу, який оптимізує застосування принципів TRIZ у складних, неповних та нечітких сценаріях і підтримує розвинені когнітивні процеси та прийняття рішень [8]. У [9] автори представили підхід до моделювання правового міркування на основі DIKWP-структури. Й. Мей та З. Гуо (Y. Duan та Z. Guo) провели системне дослідження теоретичних засад, можливості, експериментальних результатів та прикладних застосувань семантичної математики DIKWP*DIKWP для підвищення когнітивних здібностей великих мовних моделей і розвитку «штучної» свідомості [10].

Ці роботи підсилюють ідею багаторівневого представлення знання (від даних до цілі) з інтеграцією етичних та цільових компонентів. Однак, більшість з них розглядають проблеми у сфері загального штучного інтелекту або досліджують семантичні моделі для англомовного інформаційного наповнення [7, 11]. Питання узгодження об'єктивного правового контенту та суб'єктивних когніцій учасників судового процесу, як і прозорого обґрунтування прийняття судових рішень, потребує подальшого дослідження [12]. Особливо це стосується українського судочинства та українського правового контенту, оскільки українська мова вважається слабо структурованою для великих мовних моделей через складну морфологію, вільний порядок слів та недостатню кількість навчальних даних [13, 14].

Мета роботи. Метою даної роботи є розробити та обґрунтувати методологію трансформації інформаційних процесів судового розгляду з традиційних уявлень у концептуальному просторі в семантичні графічні представлення на основі моделі DIKSP (дані-інформація-знання-сенс-мета) для забезпечення ефективного узгодження україномовного правового контенту з суб'єктивними інтерпретаціями учасників судового процесу.

Для реалізації мети дослідження у роботі поставлено та вирішено такі завдання:

- формалізувати структуру DIKSP графу для представлення україномовного юридичного контенту з визначенням п'яти семантичних шарів та відображень між ними;
- розробити модель DIKSP графу сприйняття учасників судового процесу для репрезентації когнітивно-семантичного світу сторін справи;
- описати процес трансформації та алгоритм двонаправленого відображення між просторами контенту та суб'єктивних інтерпретацій (DIKSP × DIKSP);

– визначити етапи застосування семантичної моделі правового мислення від завантаження контенту до прийняття та обґрунтування рішення [15];
– продемонструвати практичне застосування розробленої методології на прикладі судової справи про корупцію з детальним поясненням міркувань на всіх рівнях DIKSP.

У цій статті запропоновано методологію трансформації інформаційних процесів судового розгляду з традиційних уявлень у концептуальному просторі в семантичні графічні представлення. Суть запропонованого підходу полягає в двонаправленому відображенні між двома модельними просторами, заснованими на основі моделі «дані–інформація–знання–сенс–мета». Один із них відображає зміст судового розгляду – DIKSP (контент): закони, факти, докази тощо. Інший представляє точку зору учасників судового процесу – DIKSP (сприйняття сторін судового процесу): їхній когнітивно-семантичний світ, включаючи цілі та інтерпретації.

Відображення між цими просторами (DIKSP × DIKSP) забезпечує двосторонній потік інформації:

– контент інтерпретується з урахуванням контексту сторін у судовому розгляді;
– міркування сторін у судовому процесі переносяться в простір правового контенту.

Для україномовного контенту такі дослідження проведено вперше.

DIKSP граф для юридичного контенту

На рис. 1 зображено схематичну структуру DIKSP графу контенту (G^{CT}) – п’ятикомпонентної моделі представлення та обробки інформації з метою аналізу судових справ.

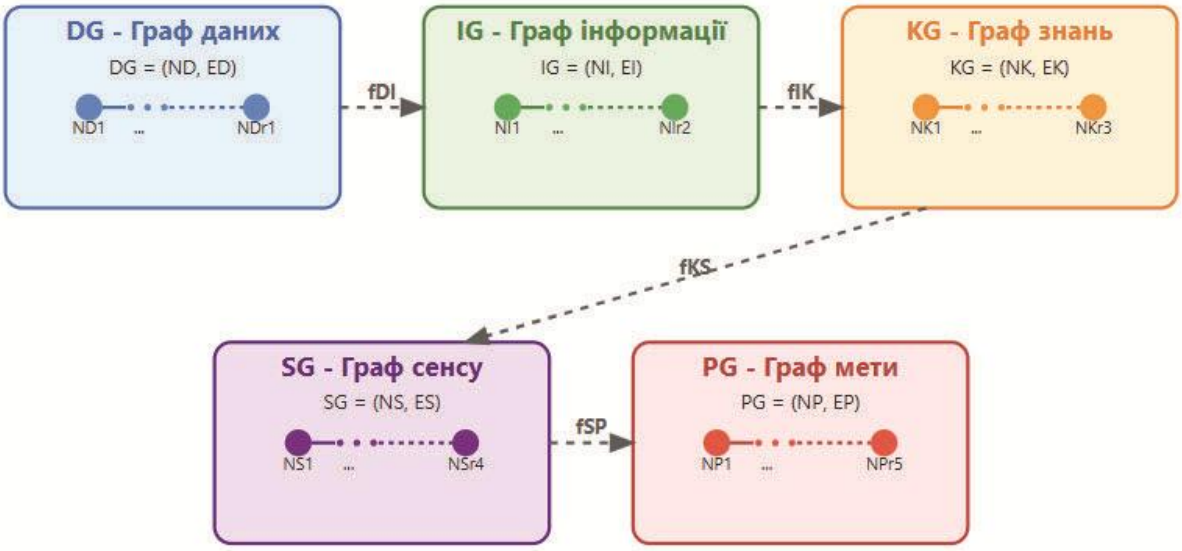


Рис. 1. Ієрархічна структура трансформації даних у цілі через інформацію, знання та сенс (мудрість)

DIKSP (контент) граф позначимо G^{CT}. Це 5-кортеж $G^{CT} = (DG, IG, KG, SG, PG)$, де:

– $DG = (N_D, E_D)$ – граф даних з вершинами N_D, які представляють базові одиниці даних зі справи (наприклад, факт внесення відомостей до Єдиного державного реєстру судових рішень [16], дата укладення договору, сума тендерної пропозиції, текстовий фрагмент статті Закону України «Про запобігання корупції») та ребра ED, які представляють зв’язки або ідентифікаційні посилання між даними (наприклад, пов’язування декларації посадової особи з її службовими повноваженнями, або зв’язування декількох документів, що стосуються одного правопорушення).

– $IG = (N_I, E_I)$ – граф інформації з вершинами N_I, які представляють інформаційні одиниці (кожна кодує значуще розрізнення або порівняння), та ребра E_I, що відображають відношення типу «відмінність», «схожість» або контекстуальні зв'язки між інформаційними одиницями. Кожен інформаційний вершина $n \in N_I$ зазвичай отримується з однієї або кількох вершин даних у N_D. Відображення f_{DI}: $N_D \to 2^{N_I}$ визначає, які інформаційні вершини походять від яких даних (наприклад, вершина даних «вартість майна у декларації – 10 млн грн» може відображатися на інформаційну вершину «вартість майна перевищує законодавчо допустимий рівень для сумісництва з посадою»).
– $KG = (N_K, E_K)$ – граф знань, де N_K включає вершини, що представляють юридичні концепції або норми (наприклад, «конфлікт інтересів», «недостовірне декларування» згідно з Законом України «Про запобігання корупції», або конкретну норму Кримінального кодексу України). E_K – ребра, що відображають логічні чи онтологічні відношення («є видом», «містить елемент», «призводить до»). Вершини знань можуть також представляти конкретизовані твердження (наприклад, «виявлено конфлікт інтересів у публічної особи» або «занижено вартість активів у декларації»), які формуються шляхом застосування загальних норм до конкретних інформаційних вершин. Відображення f_{IK}: $N_I \to 2^{N_K}$ вказує, які вершини знань активуються певними інформаційними вершинами.
– $SG = (N_S, E_S)$ – граф сенсу з вершинами N_S, які позначають конструкти вищого рівня, такі як принципи, цінності або напрацьовані практики (наприклад, «прозорість», «підзвітність», «невідворотність покарання»). Ребра E_S відображають взаємозв'язки впливу або залежності між цими принципами. Відображення f_{KS}: $N_K \to 2^{N_S}$ пов'язує вершини знань з вершинами сенсу (наприклад, концепція «недостовірне декларування» може бути пов'язана з принципом «підзвітність»).
– $PG = (N_P, E_P)$ – граф мети з вершинами N_P, що представляють цілі або наміри (наприклад, «зменшення рівня корупції в державному секторі», «забезпечення прозорості державних закупівель», «підвищення довіри громадськості»). Ребра E_P відображають ієрархію або взаємозв'язки між цілями (наприклад, «забезпечення прозорості державних закупівель» є складовою ширшої мети «зменшення рівня корупції»). Відображення f_{SP}: $N_S \to 2^{N_P}$ визначає, які цілі підтримуються певними принципами.

DIKSP граф сприйняття сторонами судового процесу

На рис. 2 представлено схематичну структуру DIKSP графу сторін судового процесу (G^{PT})

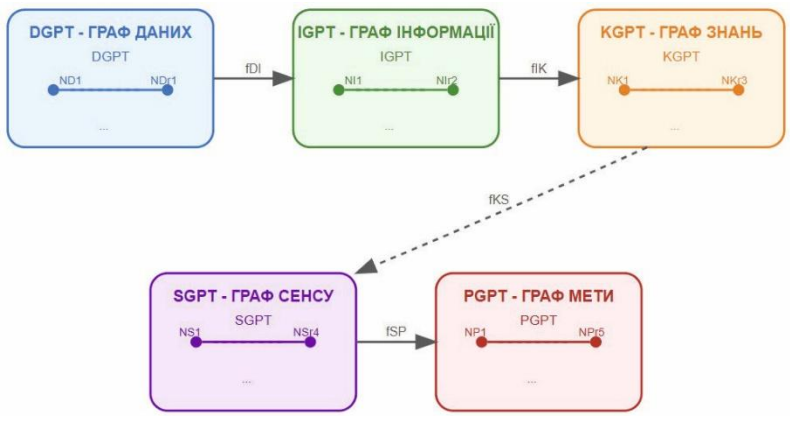


Рис. 2. Когнітивна модель учасників судового процесу: від особистих даних до процесуальних цілей

Граф DIKSP (сприйняття сторін судового процесу) позначимо G^{PT}. Це є 5-кортеж: $G^{PT} = (DG^{pt}, IG^{pt}, KG^{pt}, SG^{pt}, PG^{pt})$. Його структура аналогічна до G^{CT}, але інтерпретація вершин кожного рівня є специфічною для конкретного учасника судового процесу. Це може бути фізична особа, організація або навіть умовний "розсудливий суб'єкт" як стандарт.

Для зручності учасники судового процесу індексуються відповідно до їх приналежності до сторони судового процесу (наприклад, G^{PT_A} для сторони *A*, G^{PT_B} для сторони *B*). G^{PT} містить такі компоненти:

- $DG^{pt} = (N_D^{pt}, E_D^{pt})$, де N_D^{pt} містить набір даних, сприйнятих або наданих однією із сторін судового процесу. Це можуть бути особисті дані (наприклад, відомості про освіту чи сімейний стан) або докази з точки зору цієї сторони процесу (частково збігаються з даними контенту, частково – додаткові, відомі лише її представникам).
- $IG^{pt} = (N_I^{pt}, E_I^{pt})$, де інформаційні вершини представляють інтерпретацію або акценти відповідної сторони судового процесу. Наприклад, зауваження однієї із сторін процесу "відмінності своєї справи від типових справ" може бути представлене як інформаційна вершина. Відображення $f_{DI}^{pt}: N_D^{pt} \to 2^{N_I^{pt}}$ пов'язує такі дані з інформацією.
- $KG^{pt} = (N_K^{pt}, E_K^{pt})$, де N_K^{pt} містить знання та переконання сторони судового процесу. Сюди входить розуміння її представниками закону (правильне або хибне), знання фактів або навіть нормативні переконання ("Я діяв у рамках закону", "Суд зобов'язаний врахувати цю обставину згідно із законом"). Може включати попередній досвід ("Минулого разу в подібній ситуації я отримав попередження"). Цей рівень є когнітивною моделлю міркувань сторони у справі. Відображення $f_{IK}^{pt}: N_I^{pt} \to 2^{N_K^{pt}}$ пов'язує інформацію зі знаннями.
- $SG^{pt} = (N_S^{pt}, E_S^{pt})$ містить принципи та цінності сторони. Для фізичної особи це можуть бути уявлення про справедливість, економічну необхідність ("якщо мене позбавлять ліцензії, то я втрачу засоби до існування") або моральні міркування. Для органу це можуть бути внутрішні політики чи підходи до правозастосування ("ми надаємо пріоритет безпеці над витратами"). Відображення $f_{KS}^{pt}: N_K^{pt} \to 2^{N_S^{pt}}$ пов'язує знання з сенсом (наприклад, сторона знає, що існує певне регулювання, але вважає його застарілим і таким, що зазвичай застосовується поблажливо).
- $PG^{pt} = (N_P^{pt}, E_P^{pt})$ – цілі та призначення сторони судового процесу. Для апелянта метою, ймовірно, є "повернути ліцензію", або ширше – "досягти справедливого рішення" чи "продовжити діяльність". Можливі підцілі: відновлення репутації, мінімізація фінансових втрат тощо. Відображення $f_{SP}^{pt}: N_S^{pt} \to 2^{N_P^{pt}}$ пов'язує принципи з кінцевими цілями (наприклад, принцип справедливості у *S* може пов'язуватись із ціллю справедливого результату в *P*).

Представлена методологія визначає узгодження однієї із сторін судового процесу (апелянта) зі змістом (законом і фактами), що є критично важливим у процесі розгляду справи. Точка зору вимог закону (наприклад, законодавчий намір, дата ухвалення, обґрунтування) неявно міститься у графі змісту G^{CT}.

На рис. 3 зображено загальну схему трансформації. Спочатку юридичний контент (ліворуч) аналізується та перетворюється на граф DIKSP (контент). Цей граф містить п'ять семантичних шарів: дані, інформація, знання, сенс і мета. Аналогічно, позиція сторін процесу (праворуч) моделюється як окремий граф DIKSP (сприйняття сторін судового процесу), який також складається з таких самих шарів. Обидва графи з'єднані двонаправленими відображеннями, позначеними суцільними лініями, які пов'язують відповідні вершини. Ці зв'язки встановлюють міжшарові взаємозв'язки між контентом

та точками зору сторін судового процесу. Стрілки з позначкою «Мета» вказано на цілеспрямовані семантичні трансформації, які керують процесом інтеграції: суб'єктивна семантика сторін у справі об'єднується з об'єктивною семантикою юридичного контенту.

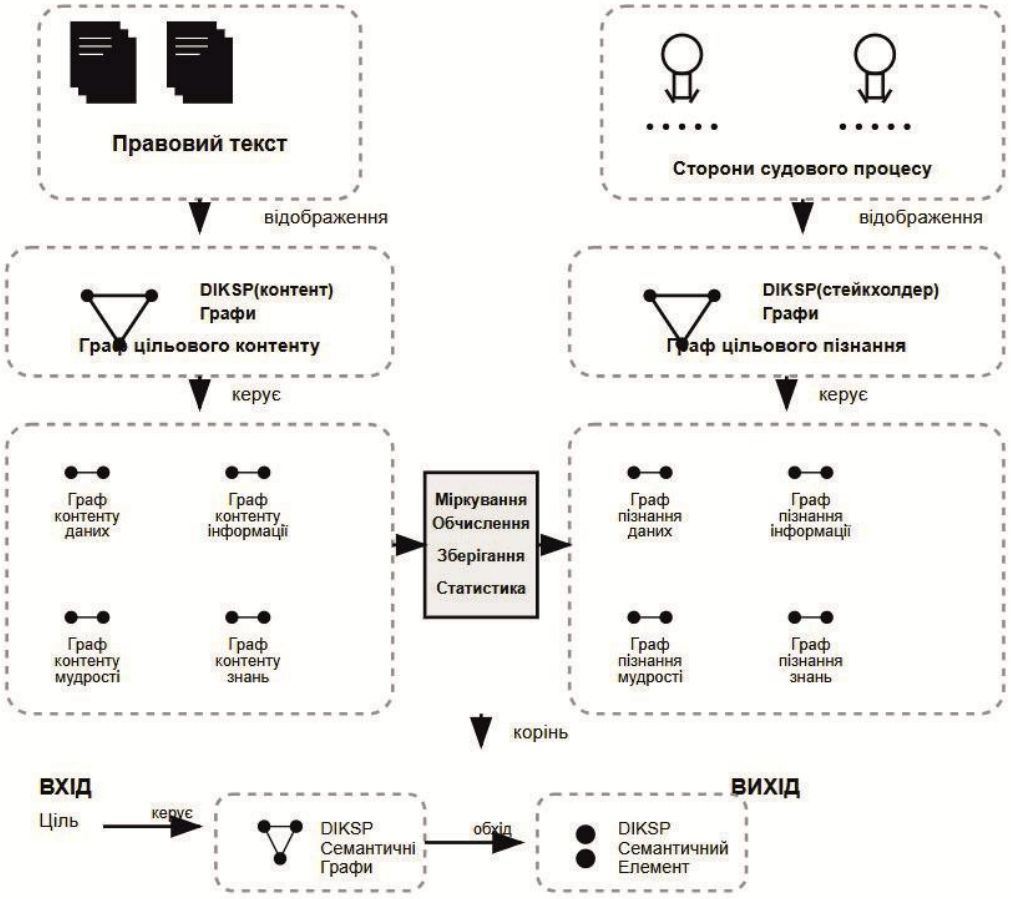


Рис. 3. Семантична модель правового мислення

При порівнянні контенту з точкою зору сторони у справі може виявитись, що учасники процесу володіють знаннями, які не відображені в контенті. У такому разі ця інформація буде додана до контенту, наприклад, у вигляді нової вершини у N_K. Або ж навпаки: якщо контент містить дані, що доповнюють знання сторони у справі, вони будуть інтегровані. Ця динаміка забезпечує, що обидва набори даних (контент і знання сторони у справі) дедалі більше відображатимуть спільне розуміння.

Етапи застосування семантичної моделі правового мислення

1. Завантаження контенту та побудова графа

Юридичний контент, зокрема закони, постанови, матеріали судових справ, докази та попередні рішення, завантажується та сегментується на шари DIKSP. Це може включати використання обробки природної мови (NLP) [17] для виділення даних та інформації (наприклад, розпізнавання іменованих сутностей для ідентифікації ключових фактів або порівнянь [5], щоб виокремити особливості справи), а також залучення юридичних баз знань для наповнення графа знань (наприклад, пов'язування ідентифікованих законів із мережею правових концепцій). Результатом цього етапу є G^{CT} (граф контенту).

Наприклад, у на цьому етапі вхідною інформацією є матеріали адміністративної корупційної справи (протоколи НАЗК, висновки експертиз, відповідні статті законів), на основі якої будуються наступні вершини:

– вершини даних для кожного релевантного факту (наприклад, дати отримання неправомірної вигоди, суми, імена фігурантів);

- вершини інформації, що відображають важливі порівняння (наприклад, «сума хабаря перевищує п’ять прожиткових мінімумів, що є значним розміром»);
- вершини знань, які кодують норми законодавства («якщо особа, уповноважена на виконання функцій держави, одержує неправомірну вигоду, її дії кваліфікуються як кримінальне правопорушення за статтею 189 КК України»);
- вершини сенсу (наприклад, принцип «забезпечення непохитної довіри суспільства до державних інститутів», виведений з антикорупційних стратегій);
- вершини мети («боротьба з корупцією»).

2. Залучення сторін судового процесу та побудова графа

Паралельно з аналізом контенту формується сприйняття сторін у справі. Наприклад, осіб, що підозрюється в корупції. Інформацію можна отримати через прямі джерела (пояснення, скарга, звернення тощо) або через когнітивну модель користувача. На основі цих даних будується граф G^{pt} (граф сприйняття сторін судового процесу).

Для цього використовуються подібні до побудови графа контенту методи (семантична екстракція, класифікація вершин), але вони адаптовані для роботи з суб’єктивною інформацією.

Наприклад, особа, підозрювана в корупції, може надати:

- вершини даних: «довідки про доходи» або «особисті обставини» («утримував малолітніх дітей»);
- вершини інформації: «наголосити, що кошти були подарунком, а не хабарем»;
- вершини знань: «посилатися на статті закону, які, на її думку, були порушені слідством»;
- вершини сенсу: «заявити про свою чесність як головний життєвий принцип» або «довести, що вчинок був вимушеним»;
- вершини мети: «уникнути кримінальної відповідальності», «зберегти репутацію», а також спільні цілі, наприклад «забезпечення законності» та «справедливого судового процесу».

3. Когнітивна узгодженість і міркування

Цей етап, проілюстрований у центрі Рис. 3, є ключовим для узгодження двох точок зору. Він реалізується через трансформації в когнітивному, концептуальному та семантичному просторах. На практиці система (або суддя) здійснює такі кроки:

- ✓ Перехід у когнітивний простір. Тут відбувається обробка даних та інформації. Система інтерпретує значення кожного фрагмента, можливо, ймовірнісно підтверджуючи певні дані як когнітивні об’єкти. На цьому етапі вирішуються розбіжності у вихідних даних. Наприклад, якщо обвинувачений заперечує певний факт, суддя приймає рішення, які дані вважати достовірними.
- ✓ Перехід у концептуальний простір. У цьому просторі організовуються відповідні юридичні поняття та норми (шар знань). Система визначає, як конкретна справа вписується в правову систему: яка норма застосовується, які визначення мають значення. Тут відбувається відображення конкретного сценарію на абстрактну структуру правових норм. На цьому етапі інтегруються знання зацікавленої сторони: якщо обвинувачений висуває аргумент щодо певної норми або посилається на іншу норму, це враховується в концептуальному просторі. Така взаємодія забезпечує узгодженість у випадку, якщо вершина знань сторони у справі відсутній у контенті. У такому разі міркування в концептуальному просторі можуть його додати. Наприклад: «обвинувачений стверджує, що слід застосувати норму Q». Якщо норма закону чинна, вона буде додана до графа знань.
- ✓ Перехід у семантичний простір. Тут розглядаються власне семантичні мережі [18], що інтегрують взаємозв’язки між даними, інформацією, знаннями, сенсом та метою [1]. Ці мережі розширюють традиційні семантичні мережі або графи знань, явно моделюючи вищі когнітивні виміри, такі як цінності та цілі. У семантичному просторі усуваються нюанси мови та контексту. Наприклад, система розуміє, що

«незначне порушення» (від зацікавленої сторони) та «порушення третього ступеня» (у контенті) позначають одне й те саме поняття, але різними словами. Таке розв'язання є ключовим для уникнення семантичних непорозумінь. Для цього можуть бути використані відображення онтологій або визначення [19].

- ✓ Узгодження через мету та міркування. Ці трансформації між просторами керовані метою. На кожному етапі система керується головними цілями (як закону, так і сторін у судовій справі), щоб узгодити розбіжності. Мета діє як евристика або ваговий коефіцієнт [21]. Якщо метою є громадська безпека (у розглянутому прикладі – запобігання корупції), то когнітивні/концептуальні невизначеності вирішуються на користь інтерпретацій, які сприяють цій меті. Виняток: якщо це суперечить цілям однієї із сторін у справі без відповідного виграшу для суспільної мети. У такому разі мета зацікавленої сторони може вплинути на іншу інтерпретацію, яка все ще мінімально задовольняє суспільну мету. Це формалізується через вплив графа мети на шлях міркування. Наприклад, якщо можна застосувати кілька норм, система надасть перевагу тій, яка краще слугує вершинам мети.

4. Прийняття та обґрунтування рішення

На заключному етапі процесу формується рішення або рекомендація. Наприклад, «визнати особу винною в корупційному правопорушенні та призначити покарання, але застосувати відстрочку виконання вироку» або «замінити кримінальну відповідальність за корупційні правопорушення на адміністративну відповідальність у вигляді штрафу». Оскільки міркування відбувалися в семантичному просторі DIKSP, можна згенерувати повне обґрунтування рішення.

Обґрунтування будується за допомогою трасування в зворотному порядку: від вершин мети (виправдання рішення) до вершин сенсу (принципів), потім до вершин знань (конкретних законів чи фактів) і, у кінцевому результаті, до суттєвих даних. Це узгодження можна представити у вигляді пояснення природною мовою:

- ✓ «Рішення про [результат] було прийнято для досягнення мети [безпека суспільства] (мета)».
- ✓ «При ухваленні цього рішення були враховані принципи [пропорційного покарання] (сенс)».
- ✓ «Згідно з Законом *Y* (знання), хоча скоєно три правопорушення (інформація з вершин даних), розміри заподіяної шкоди були незначними та обвинувачений відшкодував усі збитків (сенс), як зазначала сторона у справі. Менша санкція є достатньою для забезпечення дотримання норм і не підриває цілей громадської безпеки (узгодження мети)».

Застосування семантичної моделі до прикладу судової практики. Суддя виносить рішення у справі про корупцію: «Визнати обвинуваченого винним у корупційному злочині, передбаченому ч. 1 ст. 368 КК України, але замість позбавлення волі призначити покарання у вигляді штрафу у розмірі 1500 неоподатковуваних мінімумів доходів громадян з позбавленням права обіймати певні посади на строк 2 роки. Додатково, обвинувачений зобов’язаний пройти антикорупційну програму навчання».

Це гіпотетичне рішення спрямоване на досягнення мети закону (забезпечити невідворотність покарання та запобігти майбутній корупції), водночас дотримуючись принципу справедливості (не позбавляти волі за злочин, який не мав значних негативних наслідків), таким чином досягаючи «семантичної справедливості».

Пояснення рішення на основі DIKSP (Рис. 4).

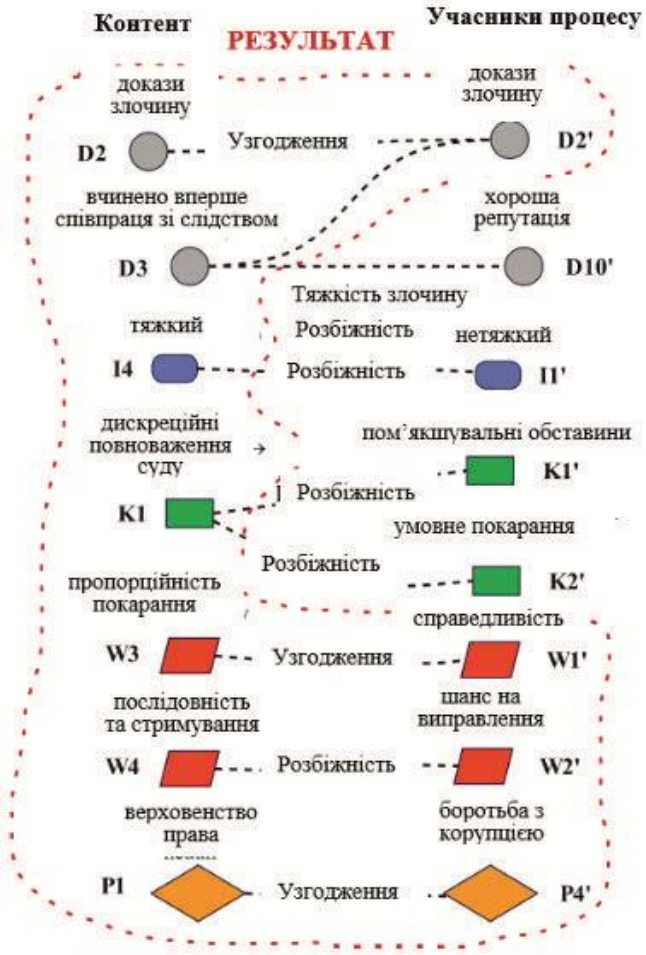


Рис. 4. Семантичний граф міркувань на основі судової справи

Мета: Рішення судді керується забезпеченням верховенства права (P1) та боротьбою з корупцією (P4'). Метою є досягнення балансу між покаранням винного та запобіганням подібним злочинам у майбутньому.

Сенс: Рішення ґрунтується на принципі пропорційності покарання (W3) та врахуванні шансу на виправлення (W2'). Хоча закон передбачає суворі заходи, суддя дійшов висновку, що в даному випадку доцільно застосувати більш м'яке покарання, яке відповідатиме тяжкості злочину та сприятиме ресоціалізації.

Знання: Згідно з K1 ("дискреційні повноваження суду"), при наявності пом'якшувальних обставин (K1') можна обрати умовне покарання або більш м'який захід (K2'). Суддя врахував, що обставини справи дозволяють винести рішення, яке узгоджується з принципом справедливості (W1') без надмірної суворості [23].

Інформація: Ключовим фактором стало визначення тяжкості злочину: злочин був класифікований як нетяжкий (I1'), що створило можливість для застосування більш гнучких підходів до покарання.

Дані: Фактичні дані включали докази злочину (D2), а також факт, що обвинувачений вчинив злочин вперше та співпрацював зі слідством (D3). Ці обставини були враховані при формуванні остаточного рішення.

Отже, рішення про призначення покарання у вигляді штрафу та додаткових обмежень «забезпечує семантичну справедливість» у справі, безпосередньо враховуючи семантику на кожному рівні: враховує дані (факт правопорушення встановлено, але обвинувачений співпрацював зі слідством); інтерпретує інформацію в контексті (порушення було серйозним з точки зору кваліфікації, але незначним за наслідками), застосовує знання правил з урахуванням нюансів (враховано дискреційні повноваження

та керівні принципи, а не лише сувору норму закону); слідує принципам сенсу (досягнення мети закону через пропорційне покарання); відповідає меті закону (протидія корупції); враховує водночас інтереси сторони у справі – обвинуваченого (можливість виправлення без ізоляції від суспільства).

Результат (рішення суду) є зрозумілим для обвинуваченого: його можна сформулювати так: «

Вас визнано винним у вчиненні корупційного правопорушення, передбаченого статтею [номер] Кримінального кодексу України. Враховуючи активне сприяння у досудовому розслідуванні та відсутність попередніх судимостей (визнання зусиль щодо спільної мети дотримання закону), призначається покарання у вигляді штрафу замість позбавлення волі відповідно до статті 69 ККУ. З урахуванням суспільної небезпеки корупційного діяння (основна мета) застосовуються додаткові заходи кримінально-правового характеру, включаючи заборону обіймати певні посади та проходження обов'язкової антикорупційної підготовки згідно із Законом України "Про запобігання корупції". Цей збалансований підхід надає обвинуваченому справедливий шанс на виправлення (дотримання принципу справедливості), водночас захищаючи суспільство від корупції (забезпечення безпеки). У випадку повторного скоєння правопорушення, вимога закону захистити суспільство зумовить більш суворе покарання (чітке пояснення того, як мета буде керувати майбутніми рішеннями).

Таке пояснення безпосередньо відображає алгоритм міркувань DIKSP, посилаючись на дані («співпраця зі слідством»), інформацію («перший злочин», «незначний розмір вигоди»), знання (правила та умови для майбутніх діянь), сенс (баланс між справедливістю та захистом суспільства) та мету (протидія корупції, надання справедливого шансу на виправлення).

Висновки. Запропоновано інноваційну методологію семантичної трансформації інформаційних процесів судових розглядів, яка забезпечує структуроване представлення як об'єктивного правового україномовного контенту, так і суб'єктивних інтерпретацій учасників судового процесу через п'ятирівневу модель DIKSP. Розроблено формальний апарат для моделювання представлення та обробки правової інформації з метою аналізу ухвалених рішень у судових справах (граф G^{CT}) та відображення когнітивних процесів учасників судового розгляду (граф G^{PT}), що включає персональні дані, інтерпретації, знання, принципи та цілі кожної сторони судового процесу. Визначено механізм двонаправленого відображення між просторами правового контенту та суб'єктивних інтерпретацій, що забезпечує взаємне доповнення обох графів та формування спільного розуміння. Виділено чотири ключові етапи застосування запропонованого підходу: побудова графу контенту, формування графу сприйняття сторін судового процесу, когнітивне узгодження через трансформації у трьох просторах (когнітивному, концептуальному, семантичному) та прийняття обґрунтованого рішення. Розроблено механізм зворотного трасування для генерування повного обґрунтування судових рішень від вершин мети до суттєвих даних, що забезпечує прозорість та зрозумілість процесу міркувань. Практичне застосування методології демонструє можливість досягнення "семантичної справедливості" через узгодження цілей закону з інтересами сторін судового розгляду на всіх рівнях DIKSP. Запропонована методологія може бути імплементована в системи підтримки прийняття судових рішень та автоматизованого обґрунтування вироків, забезпечуючи більш структурований та прозорий процес правосуддя.

Список літератури

1. Ковальчук О. Я. Організаційно-правове забезпечення цифровізації судочинства в Україні : дис. д-ра юрид. наук : 12.00.07 – адміністративне право і процес; фінансове право; інформаційне право. Київ, 2025. 573 с.

2. Васюк А. Ю. Методологічний інструментарій дослідження виконання судових рішень щодо засуджених в Україні. *Вісник Пенітенціарної асоціації України*. 2021. № 3(17). С. 76–84. URL: http://doi.org/10.34015/2523-4552.2021.3.05.
3. Дуфенюк О. М. Судові рішення у кримінальному провадженні: інформаційно-аналітичні інструменти дослідження. *Науковий вісник Ужгородського Національного Університету. Серія: Право*. 2024. Вип. 85. Ч. 4. С. 86–94. URL: https://doi.org/10.24144/2307-3322.2024.85.4.12.
4. Ковальчук О., Бабала Л., Іваницький Р. Інтелектуальна модель виявлення асоціативних правил у базах даних кримінальних правопорушень. *Вісник Хмельницького національного університету*. 2025. № 2(349). С. 188–191. URL: https://doi.org/10.31891/2307-5732-2025-349-27.
5. Kovalchuk O., Shevchuk R., Masonkova M., Banakh A. Content Analysis of Court Decisions: A GPT-4 Based Sentence-by-Sentence Data Generation and Association Rules Mining. *The First International Workshop of Young Scientists on AI for Sustainable Development* (Ternopil, Ukraine). 2024. P. 56–70.
6. Ковальчук О. Асоціативна модель підтримки прийняття рішень у кримінальному судочинстві. *Актуальні проблеми правознавства*. 2023. №3. С. 56–62. DOI: https://doi.org/10.35774/
7. Mei Y., Duan Y. The DIKWP (Data, Information, Knowledge, Wisdom, Purpose) Revolution: A New Horizon in Medical Dispute Resolution. *Appl. Sci*. 2024. Vol. 14. P. 3994.
8. Wu K., Duan Y. DIKWP-TRIZ: A Revolution on Traditional TRIZ Towards Invention for Artificial Consciousness. *Appl. Sci*. 2024. Vol. 14. P. 10865. URL: https://doi.org/10.3390/app142310865.
9. Mei Y., Duan Y. DIKWP Semantic Judicial Reasoning: A Framework for Semantic Justice in AI and Law. *Information*. 2025. Vol. 16. P. 640. URL: https://doi.org/10.3390/info16080640.
10. Duan Y., Guo Z. The DIKWP Semantic Mathematical Theory of Mathematical Subjectivization Regression. *Research Report on DIKWP*. 2025. P. 11. URL: https://doi.org/10.13140/RG.2.2.14794.48324
11. Mei Y., Duan Y., Yu L., Che H. Purpose Driven Biological Lawsuit Modeling and Analysis Based on DIKWP. Lecture Notes of the Institute for Computer Sciences. In book: *Collaborative Computing: Networking, Applications and Worksharing*. 2023. URL: https://doi.org/10.1007/978-3-031-24386-8_14.
12. Tokarieva K. S., Kovalchuk O. Ya., Kolesnikov A. P., Dzyurbel A. D., Bodnar-Petrovska O. B., Predmestnikov O. G. The Use of AI-Language Models in Judicial Proceedings: Information and Legal Aspects. *Revista Juridica Unicuritiba*. 2024. Vol. 2(78). P. 520–538. URL: https://doi.org/10.24144/2307-3322.2023.78.2.50.
13. Царьова І. Ц. Сучасний український юридичний текст: лексико-деривaційна структура : монографія. Дніпро : ЛІРА, 2020. 446 с.
14. Kovalchuk O., Banakh S., Masonkova M., Berezka K., Mokhun S., Fedchyshyn. Text Mining for the Analysis of Legal Texts. *12th International Conference "Advanced Computer Information Technologies"* (Spišská Kapitula, Slovakia, September 26–28, 2022). P. 502–505. URL: https://doi.org/10.1109/ACIT54803.2022.9913169.
15. Kovalchuk O., Shevchuk R., Chudyk N., Ivanytskyi R. Using Machine Learning Models to Decision-Making in the Justice System. *Informatyka techniczna i sztuczna inteligencja – 2024*. Wydawnictwo Naukowe Akademii Techniczno-Humanistycznej w Bielsku-Białej. P. 53–69. URL: https://doi.org/10.53052/9788367652292.03.
16. Єдиний державний реєстр судових рішень. URL: https://reyestr.court.gov.ua/.
17. Rongali S. K. Natural Language Processing (NLP) in Artificial Intelligence. *World Journal of Advanced Research and Reviews*. 2025. Vol. 25(01). P. 1931–1935. URL: https://doi.org/10.30574/wjarr.2025.25.1.0275.

18. Pereira H. B. B., Grilo M., Fadigas I. S., Souza Junior C. T., Cunha M. V., Barreto R. S. F. D., Andrade J. C., Henrique T. Systematic review of the “semantic network” definitions. *Expert Systems with Applications*. 2022. Vol. 210. P. 118455. URL: https://doi.org/10.1016/j.eswa.2022.118455.
19. Krzanowski R., Polak P. Ontology and AI Paradigms. *Proceedings*. 2022. Vol. 81. P. 119. URL: https://doi.org/10.3390/proceedings2022081119.
20. Alghamdi H., Hafeez G., Ali S., Ullah S., Khan M. I., Murawwat S., Hua L.-G. An Integrated Model of Deep Learning and Heuristic Algorithm for Load Forecasting in Smart Grid. *Mathematics*. 2023. Vol. 11(21). P. 4561. URL: https://doi.org/10.3390/math11214561.
21. Berezka K., Kovalchuk O., Banakh S., Zlyvko S., Hrechaniuk R. A Binary Logistic Regression Model for Support Decision Making in Criminal Justice. *Folia Oeconomica Stetinensia*. 2022. Vol. 22(1). P. 1–17. URL: https://sciendo.com/article/10.2478/foli-2022-0001.

COGNITIVE MODEL FOR HARMONIZING LEGAL CONTENT WITH SUBJECTIVE INTERPRETATIONS OF TRIAL PARTICIPANTS

O. Kovalchuk

West Ukrainian National University
11, Lvivska Str., Ternopil, 46009, Ukraine
Email: olhakov@gmail.com

The article proposes an innovative methodology for the semantic transformation of judicial proceedings’ information processes based on the five-level DIKSP model (Data-Information-Knowledge-Sense-Purpose). The relevance of the study is driven by the need to bridge gaps between objective legal content and subjective interpretations of trial participants, which often leads to misunderstandings and ineffective decisions. A formal apparatus has been developed for the structured representation of objective Ukrainian legal content through the G^{CT} graph and subjective interpretations of trial participants through the G^{PT} graph. Each graph contains five interconnected semantic layers: data (case facts, evidence, documents), information (significant comparisons, differences, and contextual connections), knowledge (legal norms, concepts, and legal statements), sense (principles, values, and established practices), and purpose (goals, intentions, and task hierarchies). A bidirectional mapping mechanism has been defined between legal content and subjective interpretation spaces, ensuring the mutual complementation of graphs and the formation of a shared understanding through dynamic knowledge integration. Four sequential stages of applying the proposed approach to Ukrainian judicial practice analysis have been identified: content graph construction using natural language processing; formation of trial participants’ perception graph through analysis of their positions; cognitive alignment through purposeful transformations in cognitive, conceptual, and semantic spaces; judicial decision making and justification. A backward tracing mechanism has been developed to generate complete explanations of judicial decisions, tracing from purpose vertices to factual data, thereby ensuring transparency in the reasoning process. Practical application using a corruption case example demonstrates the possibility of achieving “semantic justice” through alignment of law objectives with participants’ interests at all DIKSP levels. The methodology creates a foundation for implementation in judicial decision support systems and automated verdict justification, ensuring a more structured, transparent, and efficient justice process.

Keywords: semantic graphs, cognitive modeling, artificial intelligence, legal reasoning, legal content, cognitive-semantic space, decision justification.

DOI 10.15276/imms.v15.no4.588
УДК 004.94

Informatics and Mathematical Methods in Simulation
Vol.15 (2025), No. 4, pp. 588-596

МАТЕМАТИЧНІ МОДЕЛІ ОЦІНЮВАННЯ СТАНУ КОРИСТУВАЧА НА ОСНОВІ ЩОДЕННОЇ АКТИВНОСТІ

О. В. Корчмар, Ю. І. Бабич, М. І. Бабич

Національний університет «Одеська політехніка»
1, Шевченка пр., Одеса, 65044, Україна
Emails: akorcmar234@gmail.com, babich.u.i@op.edu.ua, babich.tiger@gmail.com

У сучасних умовах цифровізації інформаційні технології дедалі частіше застосовуються не лише для автоматизації бізнес-процесів, а й у сфері моніторингу здоров'я та психоемоційного стану людини. Одним із перспективних напрямів є побудова математичних моделей, що дозволяють перетворювати повсякденні дії користувача на кількісний показник - умовний «індекс стану». У статті розглянуто три підходи до його визначення. Перша модель (А) базується на простій сумі коефіцієнтів, що забезпечує швидку оцінку, але ігнорує контекст виконання дій. Друга модель (В) вводить вагові коефіцієнти, які змінюються залежно від часу доби, що дозволяє відобразити біоритмічні особливості організму. Третя модель (С) застосовує експоненційне затухання для опису зменшення впливу дій у часі, що забезпечує найбільш реалістичне відображення короткострокових ефектів. Для перевірки адекватності моделей розроблено експериментальну програму, яка дозволяє користувачеві створювати бібліотеку дій, задавати коефіцієнти та обирати модель для розрахунку індексу. Результати виводяться у числовій та графічній формах. Практичні тести підтвердили ефективність моделі С для відображення динамічних процесів, водночас модель А може бути корисною для швидкої оцінки, а модель В виступає проміжним варіантом. Запропонований підхід демонструє можливості поєднання класичних математичних методів із сучасними програмними засобами для створення персоналізованих систем моніторингу добробуту. Перспективним є подальше розширення моделей шляхом включення соціальних, фізіологічних та поведінкових факторів, а також використання алгоритмів машинного навчання для прогнозування змін стану.

Ключові слова: математичне моделювання, індекс стану, експоненційне затухання, вагові коефіцієнти, цифрове здоров'я, програмні системи, моніторинг добробуту.

Вступ. Сучасні інформаційні системи дедалі частіше використовуються не лише для автоматизації бізнес-процесів, управління технічними комплексами чи оптимізації виробничих операцій, а й для вирішення завдань, пов'язаних із моніторингом стану здоров'я, поведінки та психоемоційного стану людини. У час цифровізації та глобальної інформатизації суспільства все більшої актуальності набувають технології, які дозволяють обробляти значні обсяги даних про діяльність користувача та трансформувати їх у зрозумілі й корисні індикатори.

Однією з перспективних тенденцій розвитку є створення математичних моделей, які здатні узагальнювати інформацію про повсякденні дії (наприклад, сон, фізичну активність, споживання кави, перегляд відео-контенту, роботу за комп'ютером, соціальні взаємодії) і на основі цих дій визначати умовний «індекс стану». Такий індекс може інтерпретуватися, як кількісна характеристика психоемоційної чи фізичної рівноваги користувача, його продуктивності або рівня ресурсності. Важливо підкреслити, що він не є медичним діагнозом, проте може виступати як допоміжний інструмент самоспостереження та підтримки здорового способу життя.

Необхідність у подібних розробках зумовлена декількома чинниками. По-перше, сучасна людина постійно перебуває у середовищі високих інформаційних навантажень, що призводить до зростання рівня стресу та зниження концентрації. По-друге, відсутність своєчасної оцінки власного стану часто спричиняє накопичення негативних

ефектів (наприклад, виснаження, зниження когнітивних здібностей), які важко компенсувати у короткі терміни. По-третє, поява доступних мобільних пристроїв, датчиків і сенсорів відкрила нові можливості для збору персональних даних, але водночас постала проблема їх осмисленого використання.

Таким чином, актуальною науково-практичною задачею є створення системи, що поєднує математичні методи моделювання, обробку часових рядів та програмні інструменти комп'ютерних наук, дозволяючи користувачу в інтерактивному режимі оцінювати свій стан. Запропонований у цій роботі підхід базується на концепції багатомодельного розрахунку, де для одного й того ж набору дій застосовуються різні математичні алгоритми (лінійна модель, модель з урахуванням часу доби, модель експоненційного затухання). Такий підхід дає змогу порівнювати результати та вибирати найбільш адекватний метод для конкретних умов.

У статті також розглянуто програмну реалізацію створеної моделі у вигляді веб-додатку, що містить інтерфейс для введення дій, їх параметрів та візуалізації отриманого індексу у вигляді графіка чи інтерактивного кільцевого індикатора. Це робить розробку придатною для подальшого використання у навчальних цілях, у системах підтримки прийняття рішень, а також як прототип у галузі e-health та цифрового добробуту.

Огляд літератури. Аналіз сучасних досліджень свідчить про зростаючу актуальність математичного моделювання у сфері моніторингу поведінкових та фізіологічних характеристик людини. З поширенням концепцій digital health, quantified self та well-being informatics науковці шукають інструменти, які дозволяють перетворювати повсякденні дії користувача (сон, фізична активність, вживання кофеїну, перегляд відео, спілкування) у кількісні показники, що відображають його психоемоційний або фізіологічний стан.

У класичних роботах [1–3] математичне моделювання використовувалося переважно в індустріальних або технічних процесах, однак аналогічні методи почали застосовувати й у сфері людино-машинної взаємодії. Одним із ключових підходів є використання експоненційних функцій для опису процесів поступового зниження впливу дій у часі. Такий підхід дозволяє моделювати ефект «затухання». Наприклад, після чашки кави рівень бадьорості підвищується, але вже через кілька годин він повертається до базового значення. У статтях [4–5] показано, що експоненційна функція добре апроксимує подібні процеси, а зміна параметра τ (час напіврозпаду) дозволяє підлаштовувати модель під різні індивідуальні сценарії.

Другим важливим напрямом є врахування біоритмів. У працях [6–7] доведено, що ефективність дії залежить від контексту часу. Фізичні вправи зранку стимулюють продуктивність, але ввечері можуть ускладнювати засинання. Перегляд розважального контенту вночі значно сильніше впливає на когнітивні функції, ніж вдень. Для цього застосовуються вагові функції часу доби, які дозволяють гнучко масштабувати коефіцієнти моделей залежно від години виконання дії. Подібні ідеї перегукуються з класичними моделями хроно-біології та циркадних ритмів.

Третім перспективним напрямом є інтеграція математичних моделей у програмні комплекси. У публікаціях [8–9] показано, що реалізація моделей у середовищах JavaScript, Python або мобільних платформах забезпечує можливість інтерактивного розрахунку та візуалізації стану користувача. Використання інтерактивних графіків, кругових діаграм та анімованих індикаторів робить такі системи не лише науково обґрунтованими, але й зручними для повсякденного використання.

Окремий пласт літератури пов'язаний із мультифакторними моделями. Вони комбінують прості адитивні підходи (суму коефіцієнтів), вагові функції часу та експоненційні моделі з методами машинного навчання для прогнозування майбутніх станів. У статтях [2, 7, 10] описано приклади систем, де до математичних моделей додаються алгоритми класифікації і регресії, що підвищує точність і дозволяє індивідуалізувати результат. Таким чином, сучасна тенденція полягає не у виборі однієї

«універсальної» формули, а в побудові комплексних моделей, які поєднують різні підходи та враховують специфіку конкретного користувача.

Отже, сучасний науковий дискурс свідчить, що моделювання психоемоційного та фізичного стану користувача поступово переходить від вузько-спеціалізованих досліджень до практичних програмних реалізацій. Запропонована у даній роботі система базується саме на цій ідеї, поєднати класичні математичні методи (сума коефіцієнтів, вагові функції, експоненційне затухання) з сучасними засобами веб-програмування та інтерактивної візуалізації, створивши доступний інструмент для самоспостереження та аналізу.

Мета роботи. Метою даної роботи є розробка та дослідження математичних моделей, що дозволяють обчислювати умовний індекс стану користувача на основі його повсякденної активності. Під індексом стану розуміється узагальнений показник, який відображає сукупний вплив різних факторів - від відпочинку та фізичної активності до споживання стимулюючих напоїв чи взаємодії з цифровим контентом. Особливістю дослідження є поєднання простих та інтуїтивно зрозумілих методів з більш складними математичними підходами, що враховують часову структуру та динаміку впливів.

Для досягнення поставленої мети запропоновано три моделі. Перша модель ґрунтується на простій сумі коефіцієнтів, що характеризують позитивний або негативний вплив окремих дій. Вона є найпростішою для сприйняття та може використовуватися як базова. Друга модель вводить часову залежність, враховуючи, що ефективність тієї самої дії може змінюватися залежно від періоду доби. Наприклад, сон у нічний час має інший ефект, ніж короткий денний відпочинок. Третя модель відображає процес поступового затухання впливу з часом за допомогою експоненційних функцій, що наближає її до реальних фізіологічних процесів та забезпечує більш гнучкий опис накопичених дій.

Розробка та аналіз цих моделей дозволяє створити адаптивну систему, яка забезпечує користувачеві інструмент для відстеження власного стану у динаміці. Таке рішення може застосовуватися як у сфері особистого моніторингу здоров’я, так і у більш широких контекстах - від досліджень впливу стилю життя на продуктивність до інтеграції в мобільні додатки, орієнтовані на підтримку психоемоційного балансу.

Основний розділ. У межах дослідження було запропоновано три математичні моделі, які відрізняються рівнем складності та глибиною врахування додаткових факторів. Вибір цих моделей зумовлений необхідністю як забезпечення простоти розрахунків для користувача, так і точності у відображенні динаміки змін індексу стану.

Перша модель (А) є базовою і передбачає використання принципу адитивності. Вона спирається на просту суму коефіцієнтів, які характеризують внесок кожної окремої дії у формування інтегрального показника. Формула має вигляд (1):

$$I_A = \sum a_i, \tag{1}$$

де a_i - внесок і-ої дії.

У даному випадку кожна дія розглядається як рівнозначний елемент системи, а її вплив не залежить від зовнішніх чинників. Такий підхід дозволяє отримати швидку оцінку, що може бути використана для початкового моніторингу або як спрощена метрика. Проте недоліком цього підходу є відсутність урахування контексту виконання дій. Наприклад, споживання кави вранці та ввечері впливатиме на організм по-різному, але в моделі А ці впливи будуть оцінені однаково. Аналогічна ситуація може спостерігатися й у випадку фізичних вправ. Ранкова пробіжка сприяє підвищенню працездатності протягом дня, тоді як вечірнє тренування, особливо інтенсивне, може ускладнити засинання та призвести до зниження якості сну. Таким чином, модель А надає базову кількісну оцінку, але не враховує індивідуальні особливості та зовнішні умови, що обмежує її застосування у більш складних сценаріях моделювання.

Друга модель (B) враховує часову залежність, що є ключовим аспектом у дослідженні добових ритмів людини. Відомо, що фізіологічні процеси організму змінюються протягом доби, і саме тому ефект від одних і тих самих дій у різний час може значно відрізнятися. Для моделювання цього явища вводиться вагова функція $w(t_i)$, яка залежить від моменту часу t_i, коли дія була виконана. Розрахунок здійснюється за формулою (2):

$$I_B = \sum (b_i * w(t_i)) , \quad (2)$$

де b_i - коефіцієнт, що описує інтенсивність впливу дії, $w(t_i)$ -ваговий множник, що змінюється залежно від часу доби.

У моделі B кожна дія описується не лише власним коефіцієнтом інтенсивності b_i але й додатковим ваговим множником $w(t_i)$ що залежить від часу доби. Це дозволяє врахувати біологічні ритми та добову циклічність організму. Відомо, що фізіологічні процеси людини мають циркадну природу, змінюється рівень гормонів, температура тіла, активність нервової системи. Тому одна й та сама дія може мати суттєво різний ефект залежно від того, коли вона виконана.

Наприклад, фізичне навантаження, виконане вранці, часто сприяє активації організму, покращує концентрацію уваги та підвищує працездатність у першій половині дня. У моделі B це відображається через ваговий коефіцієнт, який у ранкові години збільшує значущість дії. У свою чергу, аналогічне навантаження пізно ввечері може призвести до збудження нервової системи, утруднення процесу засинання й навіть до зниження якості нічного відпочинку. Для цього випадку ваговий множник має зменшене значення або навіть може змінювати знак впливу, якщо дія визнана шкідливою у даному часовому контексті.

Аналогічно працює і для інших дій. Вживання кави у ранкові години може позитивно впливати на концентрацію уваги та енергійність, але після 18:00 ця ж дія потенційно погіршує якість сну, тому ваговий коефіцієнт у моделі B буде знижуватися. Те саме стосується соціальних активностей. денний контакт з іншими людьми може підвищувати емоційний тонус, але надмірна активність у нічний час здатна стати джерелом перевтоми.

Таким чином, модель B можна розглядати як більш гнучку у порівнянні з моделлю A, адже вона дозволяє враховувати контекст виконання дії. Вона відображає залежність ефективності або шкідливості активності від часу, наближаючи математичну модель до реальних умов функціонування людського організму. У практичному застосуванні це означає, що система може рекомендувати не лише набір дій, але й оптимальний час для їх виконання.

Третя модель (C) спрямована на формалізацію процесу зменшення впливу дій з часом. У реальному житті більшість подій мають тимчасовий ефект, який не є сталим. Наприклад, випита чашка кави бадьорить лише протягом кількох годин, після чого її стимулюючий вплив зникає. Аналогічно, фізичні вправи підвищують рівень енергії та покращують настрій, однак цей ефект триває певний проміжок часу, після чого організм повертається до базового стану. Навіть негативні фактори, такі як стрес або конфлікт, мають схильність до поступового зменшення інтенсивності впливу, якщо не підкріплюються новими подіями.

Для опису цього явища в моделі використовується експоненційна функція затухання, що є універсальним інструментом у багатьох наукових дисциплінах. Використання експоненти дозволяє передати природний закон спадання впливу, коли значення поступово зменшується, але ніколи повністю не зникає.

Математично модель C можна записати у такому вигляді (3):

$$I_C = \sum (c_i * e^{\frac{-\Delta t_i}{\tau}}) , \quad (3)$$

де c_i - коефіцієнт, що характеризує дію, Δt_i - час, що минув після виконання дії, τ - параметр, який задає швидкість зменшення впливу (аналог періоду напіврозпаду).

Якщо τ обрати невеликим, ефекти дій швидко зникають, і індекс буде чутливим до нещодавніх подій. Якщо ж τ великий, тоді минулі події довше зберігають свій внесок у систему, і-індекс відображатиме більш інерційний стан. Це дозволяє гнучко налаштовувати модель залежно від поставлених завдань, чи необхідно фіксувати миттєві зміни, чи важливо враховувати довготривалі ефекти.

У практичному застосуванні експоненційна модель дуже корисна для відстеження динаміки стану користувача впродовж доби. Наприклад, після тренування на ранковому етапі спостерігається високий внесок у індекс, але з часом він плавно зменшується. Або ж негативний вплив стресової події, що стався вдень, у вечірні години ще зберігає частину своєї сили, проте поступово втрачає актуальність.

Таким чином, модель С є найбільш реалістичною серед запропонованих, оскільки враховує природну динаміку затухання ефектів. Порівняння трьох моделей демонструє поступове ускладнення підходів, від простої статичної суми (модель А), до врахування добових ритмів (модель В). До моделі з експоненційним спаданням впливу (модель С), яка найбільш адекватно відображає поведінку системи в реальному часі. Така багаторівнева структура дозволяє адаптувати інструмент як для швидкого аналізу, так і для точного прогнозування стану користувача.

Практичні результати та їх аналіз. Всі три математичні моделі (А, В та С) було реалізовано в експериментальному середовищі, яке дає змогу перевіряти їхню роботу на прикладах повсякденних дій користувачів. Кожна модель має свої переваги та обмеження, це дозволяє розглядати їх у якості різних рівнів складності під час моніторингу стану користувача.

Практична апробація показала, що модель А забезпечує найшвидший результат і може бути використана у випадках, де необхідна груба оцінка стану або базова метрика для початкового аналізу. Вона є зручною у випадках, коли дії мають однакову вагу, а часовий контекст не відіграє ключової ролі.

Модель В, що враховує вагові коефіцієнти залежно від часу доби, дозволила виявити цікаві закономірності. Наприклад, фізична активність у ранкові години мала позитивний вплив на індекс, тоді як аналогічна активність у пізній вечір знижувала його. Це підтверджує важливість урахування циркадних ритмів у математичних моделях подібного типу. Такий підхід приближує результати до реальних біологічних процесів, роблячи оцінку більш адаптивною та персоналізованою.

Модель С з експоненційним затуханням виявилася найбільш адекватною для відображення динаміки короткострокових ефектів. Було встановлено, що дії з тимчасовим впливом (кава, медитація, перегляд відео, короткий сон) найбільш коректно описуються саме через поступове зменшення їхнього внеску в загальний індекс. Результати моделі С найбільше відповідають суб’єктивним відчуттям.

Окрім теоретичної частини, було розроблено демонстраційний інструмент, у якому користувачі можуть, додавати власні дії та призначати їм коефіцієнти за шкалою та обирати одну з трьох моделей для розрахунку індексу.

На рис. 1 показано інтерфейс для додавання власної дії, де користувач задає її параметри та коефіцієнти для моделей А, В, С.

На рис. 2 наведено приклад відображення індексу стану у вигляді кільцевої діаграми, яка оновлюється залежно від обраної моделі та дозволяє швидко оцінити показник.

На рис. 3 представлено графік динаміки індексу стану (умовні дані) для трьох моделей. Видно, що модель А забезпечує стабільність, модель В враховує добові ритми, а модель С описує затухання ефектів у часі.

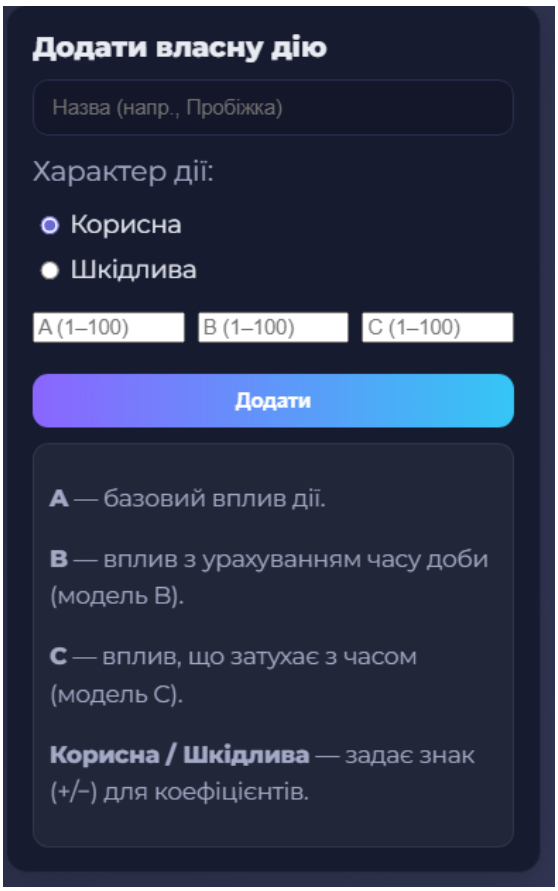


Рис.1. Форма для додавання власної дії

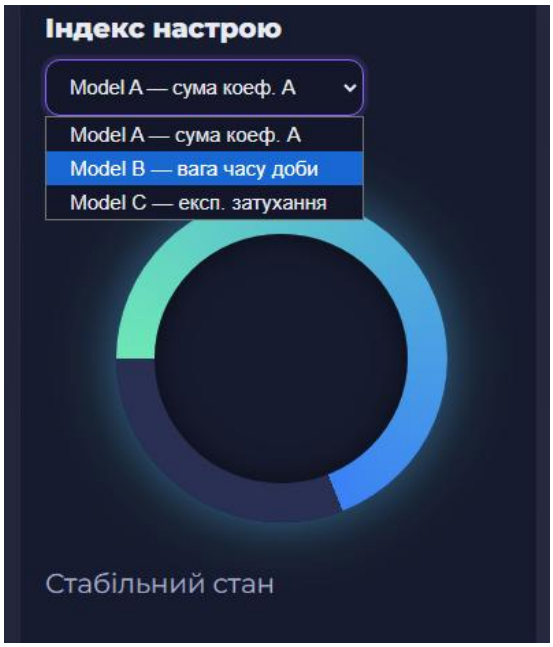


Рис.2. Кільцева діаграма індексу стану користувача

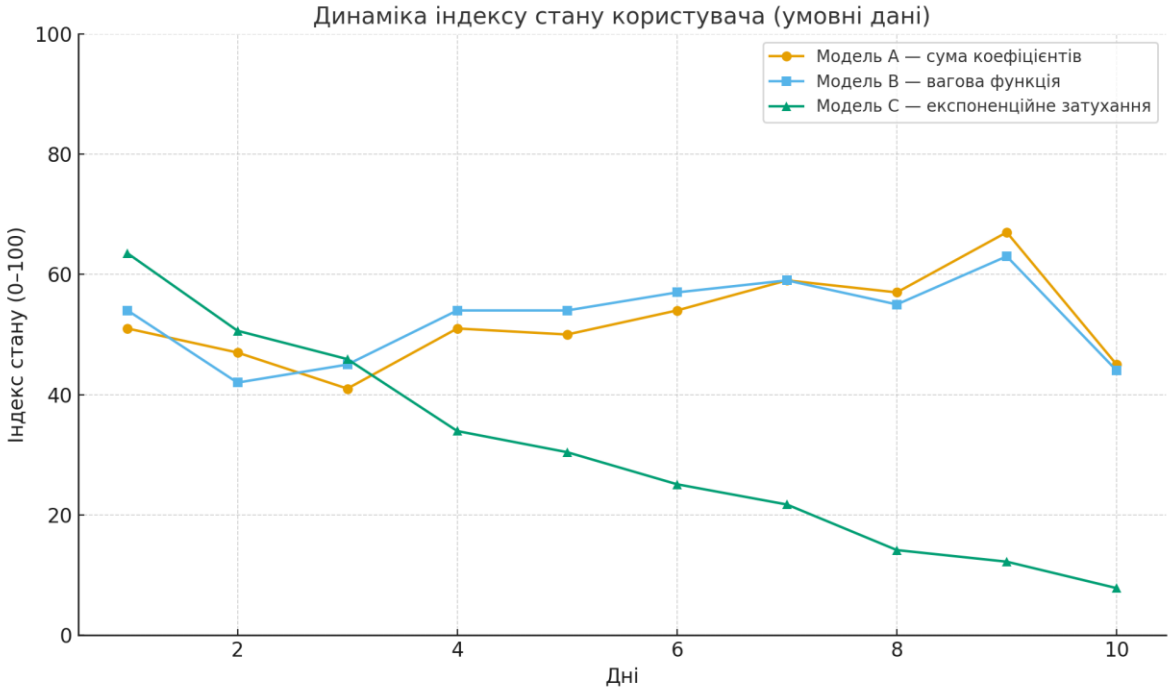


Рис. 3. Графік динаміки індексу для моделей A, B, C

Ця реалізація продемонструвала, що математичні формули можуть бути безпосередньо інтегровані у програмні комплекси й візуалізовані через сучасні інтерфейси. Особливу увагу було приділено тому, щоб результат обчислень подавався користувачу у зручній та зрозумілій формі - індикатори, графіки, повідомлення з порадами.

У перспективі подібні інструменти можуть бути інтегровані з пристроями (фітнес-трекери, смарт-годинники), що забезпечить автоматичний збір даних про активність, сон або пульс. Це дозволить не лише перевірити точність моделей, а й удосконалити їх завдяки кореляції з об’єктивними біометричними показниками.

Крім того, запропоновані моделі можуть бути використані як основа для машинного навчання, алгоритми зможуть прогнозувати майбутні зміни індексу на основі історії дій користувача. Наприклад, система зможе попередити, що надмірне споживання кофеїну після 18:00 з високою ймовірністю негативно позначиться на стані наступного дня.

Таким чином, отримані результати свідчать про те, що запропоновані формули мають універсальний характер. Вони придатні для застосування в освіті, медицині, бізнесі та повсякденному житті. Демонстраційна реалізація підтвердила практичну цінність підходу й водночас окреслила можливості для подальших досліджень, зокрема інтеграції з багатофакторними системами оцінювання, підключення до сенсорних пристроїв та використання у прогностичних моделях.

Висновки. У роботі було запропоновано три математичні моделі для оцінювання умовного індексу стану користувача, кожна з яких відображає різний рівень складності та деталізації. Найпростіший варіант - модель A - ґрунтується на звичайному підсумовуванні коефіцієнтів дій. Вона забезпечує швидку та базову оцінку, що може використовуватися для початкового моніторингу, проте її обмеженням є відсутність урахування контекстних чинників, зокрема часу чи тривалості впливу.

Більш гнучким є підхід моделі B, де передбачено використання вагових коефіцієнтів залежно від часу доби. Завдяки цьому система враховує природні біоритми людини, а результати стають ближчими до реальних умов. Так, дії, виконані у ранковий чи вечірній час, можуть мати різний ефект на організм, і математична модель здатна це коректно відобразити.

Найбільш реалістичною та наближеною до природних процесів виявилася модель C, яка описує динаміку поступового згасання ефектів у часі. Використання експоненційної функції дозволяє змоделювати, як вплив певної дії зменшується. Чашка кави підвищує бадьорість лише протягом кількох годин, фізичні вправи забезпечують покращення настрою на певний проміжок часу, тоді як стресові ситуації поступово втрачають інтенсивність. Такий підхід робить результати більш стійкими та адекватними до умов реального життя.

В рамках програмної реалізації застосунку буде враховано практичну значущість обраних підходів. Така система дозволить користувачеві вводити власні дії та одразу спостерігати за зміною індексу стану у різних моделях. Це зробить програму корисним навчальним інструментом для студентів і дослідників, а також створить основу для подальших удосконалень. У перспективі запропоновані математичні моделі можуть бути інтегровані у більш складні системи моніторингу, наприклад, у поєднанні з пристроями або алгоритмами машинного навчання, що відкриває можливості для їх використання у медицині, спорті, освітніх практиках та побутових застосуваннях.

Список літератури

1. Pereira C., Aguilar P., Saboia I., Barreto I., Theophilo R. Systematic mapping of digital health apps – A methodological proposal based on the World Health Organization classification of interventions. *Digital Health*. 2022. Vol. 8. DOI:10.1177/20552076221129071.
2. Leuzzi G., Job M., Scafoglieri A., Testa M. Smartphone Apps and Wearables for Health Parameters in Young Adulthood: Cross-Sectional Study. *JMIR Human Factors*. 2025. Vol. 12.
3. Liu, S. A "No-Code" App Design Platform for Mobile Health Research: Development and Usability Study. *JMIR Formative Research*. 2022. Vol. 6, № 8. P.e38737.
4. Developing a Cross-Platform Application for Integrating Real-Time Time-Series Data from Multiple Wearable Sensors. *Eng. Proc*. 2023. Vol. 58(1), article 4. DOI:10.3390/ecsa-10-16185.
5. Quality, Usability, and Effectiveness of mHealth Apps and Their Role of Artificial Intelligence: Current Scenario and Challenges. *JMIR*. 2023. Vol.25. P. e44030.
6. Хаханова А., Абдулаєв В. Цифрове моделювання соціальних процесів. Харківський національний університет радіоелектроніки. *CDS*. 2023. Вип. 5. № 1. С. P.47-56. DOI:10.23939/cds2023.01.047.
7. Корчинський І.О., Фірман Н.А. Цифрова медицина: особливості та проблеми становлення в Україні. *Цифрова економіка та економічна безпека*. 2020-2021.
8. Орєхов О.С., Фаріонова Т.А. Математичні моделі для оцінювання розміру Java-застосунків. *Вісник Херсонського національного технічного університету*. 2024. № 2.
9. Martynova O., Clarification of the Theoretical Foundations of Modeling the Assessment of Enterprise Activity Using a Balanced Scorecard. *Математичні методи, моделі та інформаційні технології в економіці*. 2023. № 1(88).
10. Huda O., Kyrylov S., Kyrylova L. Математичні системи для реалізації штучних нейронних мереж, орієнтованих на хмарові обчислення. *Системні технології*. 2023. № 6. С. 149-155. DOI:10.34185/1562-9945-6-149-2023-06.

О. В. Корчмар, Ю. І. Бабич, М. І. Бабич

MATHEMATICAL MODELS FOR EVALUATING USER STATE BASED ON DAILY ACTIVITY

O.V. Korchmar, Y.I. Babych, M.I. Babych

National Odesa Polytechnic University
1, Shevchenko Ave., Odesa, 65044, Ukraine

E-mail: akorcmar234@gmail.com, babich.u.i@op.edu.ua, babich.tiger@gmail.com

In the modern conditions of digitalization, information technologies are increasingly applied not only for the automation of business processes but also in the field of health and psycho-emotional state monitoring. One of the promising directions is the construction of mathematical models that allow transforming the user's everyday actions into a quantitative indicator - a conditional "state index." The article examines three approaches to its calculation. The first model (A) is based on a simple sum of coefficients, which provides a quick assessment but ignores the context of actions. The second model (B) introduces weighting coefficients that vary depending on the time of day, thus reflecting the circadian characteristics of the human body. The third model (C) applies exponential decay to describe the reduction of action influence over time, which ensures the most realistic representation of short-term effects. To verify the adequacy of the models, an experimental software system was developed, allowing the user to create a library of actions, assign coefficients, and choose a model for index calculation. The results are presented in both numerical and graphical forms. Practical tests confirmed the effectiveness of Model C in representing dynamic processes, while Model A can be useful for quick evaluations, and Model B serves as an intermediate option. The proposed approach demonstrates the potential of combining classical mathematical methods with modern software tools to create personalized well-being monitoring systems. A promising direction for further research is the extension of models by including social, physiological, and behavioral factors, as well as the use of machine learning algorithms for state change prediction.

Keywords: mathematical modeling, state index, exponential decay, weighting coefficients, digital health, software systems, well-being monitoring.

DOI 10.15276/imms.v15.no4.597
УДК 631.559:004.942

Informatics and Mathematical Methods in Simulation
Vol.15 (2025), No. 4, pp. 597-607

ДВОКОМПОНЕНТНА АДАПТИВНА МОДЕЛЬ ДИНАМІКИ ВЕГЕТАЦІЙНИХ ІНДЕКСІВ ДЛЯ ПРОГНОЗУВАННЯ УРОЖАЙНОСТІ СІЛЬСЬКОГОСПОДАРСЬКИХ КУЛЬТУР

М. В. Мачуляк

Західноукраїнський національний університет
11, Львівська, Тернопіль, 46009, Україна
Email: Mvmach9@gmail.com

Запропоновано двокомпонентну адаптивну модель динаміки вегетаційних індексів, яка на відміну від відомих містить дискретну адаптивну компоненту на основі комбінації поточних та історичних даних та неперервну апроксимаційну компоненту на базі системи диференціальних рівнянь Моно. Розроблено метод ідентифікації дворівневої адаптивної моделі урожайності, який використовує пояснюючі змінні у вигляді моделей вегетаційних індексів NDVI та MTCI замість їх безпосередніх значень. Запропонована архітектура геоінформаційної системи на базі PostGIS забезпечує ефективну інтеграцію просторових даних від БПЛА та LiDAR систем. Математична модель виявлення ущільнених ділянок ґрунту на основі рівнянь Моно показала підвищення точності у півтора рази порівняно з лінійними методами. Експериментальна апробація на даних пшениці та рису підтвердила ефективність підходу з середньою похибкою прогнозування урожайності п'ятькома з трьома відсотка для ансамблевих методів. Отримані результати демонструють перспективність застосування запропонованого підходу для розвитку систем точного землеробства та підтримки прийняття агрономічних рішень.

Ключові слова: вегетаційні індекси, модель Моно, прогнозування урожайності, адаптивне моделювання, дистанційне зондування, геоінформаційні системи, точне землеробство, підтримка прийняття рішень, екологічний моніторинг

Вступ. Забезпечення продовольчої безпеки та підвищення ефективності агровиробництва є критично важливими завданнями сучасності. Традиційні методи оцінки урожайності базуються на статистичних даних регіонального або національного рівня, що не забезпечує необхідної просторової та часової роздільності для прийняття оперативних управлінських рішень [1]. Сучасні технології точного землеробства, зокрема використання безпілотних літальних апаратів (БПЛА) у поєднанні з геоінформаційними системами (ГІС), дозволяють отримувати високоточну інформацію про стан посівів через вегетаційні індекси. Нормалізований різницевий вегетаційний індекс (NDVI) та індекс хлорофілу наземної рослинності (MTCI) є надійними індикаторами фотосинтетично активної біомаси та вмісту хлорофілу відповідно [2]. Однак існуючі підходи до моделювання динаміки вегетаційних індексів мають суттєві обмеження: статистичні моделі не враховують нелінійну природу біологічних процесів, а моделі машинного навчання вимагають великих обсягів тренувальних даних [3].

Мета дослідження - розробка двокомпонентної адаптивної моделі динаміки вегетаційних індексів та методу ідентифікації дворівневої моделі урожайності для підвищення точності прогнозування в умовах обмежених даних.

Теоретичні основи дослідження. Сучасні підходи до моделювання урожайності можна класифікувати на три основні групи: статистичні, імітаційні та методи машинного навчання [5]. Статистичні моделі базуються на емпіричних залежностях між вегетаційними індексами та урожайністю. Baez-Gonzalez et al. [6] використовували множинну лінійну регресію для зв'язку NDVI з урожайністю кукурудзи, отримавши коефіцієнт детермінації $R^2 = 0.65$. Однак такі моделі не враховують динамічну природу росту рослин. Імітаційні моделі, такі як DSSAT та APSIM, моделюють біофізичні

процеси росту, але вимагають великої кількості параметрів і складні для калібрування [7]. Моделі машинного навчання демонструють високу точність, але мають обмежені пояснювальні властивості. Zhang et al. [8] використовували згорткові нейронні мережі для прогнозування урожайності пшениці з точністю $RMSE = 0.12$ т/га.

Перспективним напрямком є використання диференціальних рівнянь для моделювання біологічних процесів. Модель Моно, спочатку розроблена для опису росту мікроорганізмів, успішно адаптується для моделювання росту вищих рослин [9].

Постановка проблеми. Побудова ГІС на базі PostGIS є оптимальним рішенням для інтеграції просторових та атрибутивних даних. PostGIS забезпечує понад 400 просторових функцій, підтримку растрових даних та ефективну індексацію за допомогою GiST індексів. Структура основних відношень ГІС базується на центральній просторовій сутності – полі (farm_fields). Ключові компоненти включають:

- farm_fields: межі полів (field_id, name, area_ha, field_geometry);
- field_blocks: технологічні ділянки (block_id, field_id, block_geometry);
- yield_data: щорічні дані урожайності (yield_id, field_id, year, crop_type, yield_t_ha);
- vegetation_indices: динаміка вегетаційних індексів (veg_id, field_id, capture_date, ndvi_mean, source)

Для зберігання растрових даних використовується гібридний підхід: метадані зберігаються в реляційних таблицях (*canopy_height_metadata*, *dtm*), тоді як самі растрові файли розміщуються у файловій системі або хмарному сховищі. Така архітектура забезпечує ефективний доступ до просторових характеристик рослинного покриву за даними LiDAR.

Відомо, що ущільнення ґрунту призводить до зниження врожайності на 15–30 %**,** що обумовлює необхідність своєчасного виявлення ущільнених ділянок. Для цього у роботі запропоновано використовувати залежність між висотою рослин $X(t)$ та опором проникненню ґрунту, формалізовану на основі моделі Моно [4].

Аналіз експериментальних даних показує S-подібний характер залежності, який адекватно описується системою диференціальних рівнянь Моно:

$$\begin{cases} \frac{dX(t)}{dt} = p_1 \frac{X(t)S(t)}{p_2+S(t)} \\ \frac{dS(t)}{dt} = -p_3 \frac{X(t)S(t)}{p_2+S(t)} \end{cases} \quad (1)$$

де: $X(t)-$ висота рослин; $S(t)-$ потенціал росту; t – час; p_1, p_2, p_3 – параметри моделі, що визначають інтенсивність росту та споживання ресурсу.

Для зменшення впливу шумів, зумовлених похибками фотограмметрії та лазерного сканування, експериментальні дані попередньо згладжувалися гаусівським фільтром, який задається операцією згортки [6]:

$$Y[i] = \sum_{j=-k}^{k} D[i-j]K[j] \quad (2)$$

де: D – вихідні дані; K – ядро Гаусса; k – радіус ядра згладжування.

У випадку одновимірного згладжування ядро фільтру відповідає щільності нормального розподілу із нульовим математичним сподіванням:

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \quad (3)$$

де: x – відстань від центру ядра; σ – стандартне відхилення, ступінь згладжування даних.

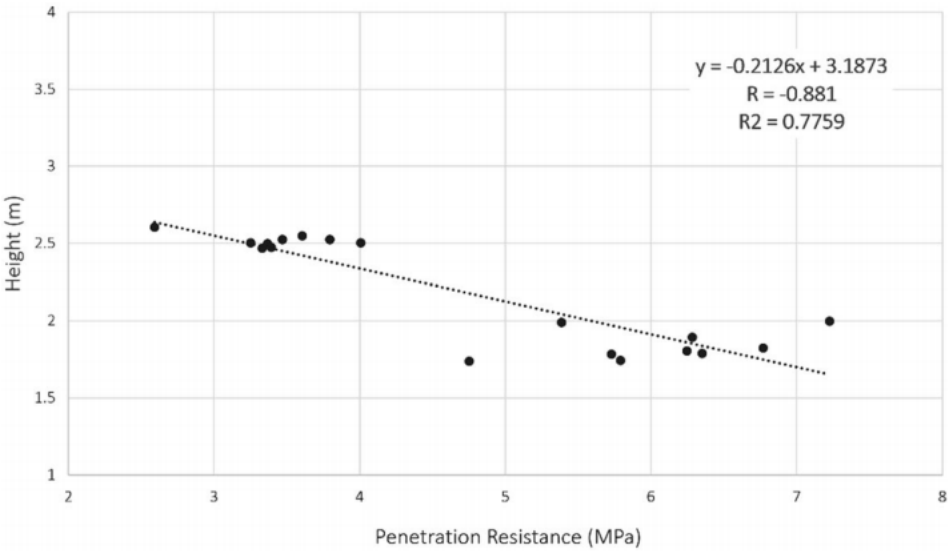


Рис. 1. Графік лінійної регресії між висотою та опором проникненню ґрунту

Після згладжування та нормалізації даних було виконано регресійний аналіз, який дозволяє розглядати локальну ділянку S-подібної кривої моделі Моно як квазілійну залежність між висотою рослин і опором проникненню ґрунту. Результати такої апроксимації наведено на рисунку 1, де показано лінійну регресію між висотою рослин та опором проникненню ґрунту, що підтверджує негативний вплив ущільнення на параметри росту рослин.

Для підвищення достовірності подальшого моделювання було виконано попередню обробку експериментальних даних**,** яка включала видалення викидів, згладжування часових рядів гаусівським фільтром та нормалізацію висотних показників, отриманих за даними LiDAR. Результати зазначених етапів попередньої обробки наведено на рис. 2, де показано зменшення шумової складової та вирівнювання тренду висоти рослин.

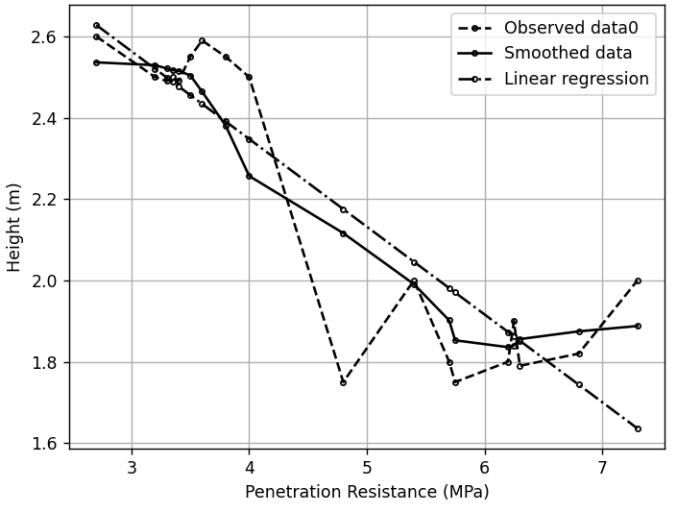


Рис. 2. Результати попередньої обробки даних

Для врахування реальних значень висоти та усунення впливу локального мікрорельєфу застосовувалося лінійне перетворення:

$$X = X^r - X^{min} \tag{4}$$

де: X^r – виміряна висота рослин за даними LiDAR; X^{min}– мінімальне значення висоти в межах досліджуваної ділянки; X – приведене значення висоти, що використовується в моделі.

Результати обговорення. Застосування перетворення (4) дозволило привести експериментальні дані до єдиної шкали та забезпечити коректну ідентифікацію параметрів моделі Моно. На основі нормалізованих даних було побудовано модель Моно співвідношення між ущільненістю ґрунту та висотою рослин, яка відображає нелінійний характер впливу ущільнення на ріст рослин.

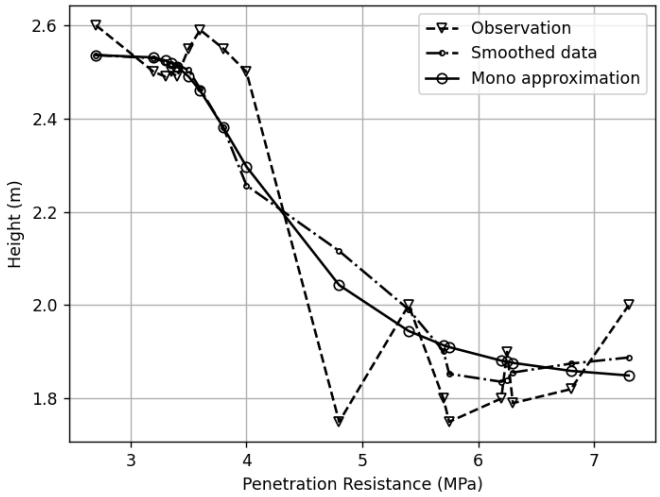


Рис. 3. Модель Моно співвідношення між ущільненістю ґрунту та висотою рослин

Порівняльний аналіз точності апроксимації показав, що модель Моно має вищу прогностичну здатність у порівнянні з лінійною регресією. Зокрема, для лінійної моделі максимальна відносна похибка становить 16,3 %**,** а середня — 4,5 %**,** тоді як для моделі Моно відповідні значення зменшуються до 10,4 % та 3,3 %**.** Це підтверджує доцільність використання нелінійної моделі Моно для аналізу впливу ущільнення ґрунту на параметри росту рослин. Для валідації запропонованого підходу використовувалися дані динаміки NDVI по 15 полях пшениці з північної Бельгії протягом трирічного періоду (2018-2020).

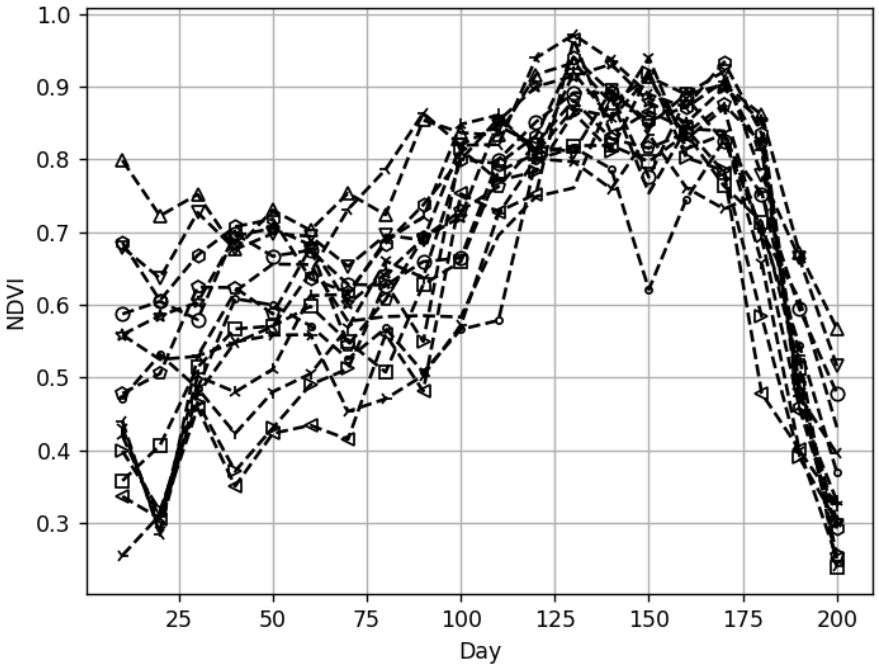


Рис. 4. Динаміка вегетаційного індексу NDVI

Перехід до кумулятивних значень NDVI забезпечив стабільність процесу моделювання та усунення випадкових коливань вихідних даних. Кумулятивні значення

формували щільний пучок траєкторій, що свідчить про узгодженість динаміки розвитку рослин на різних полях.

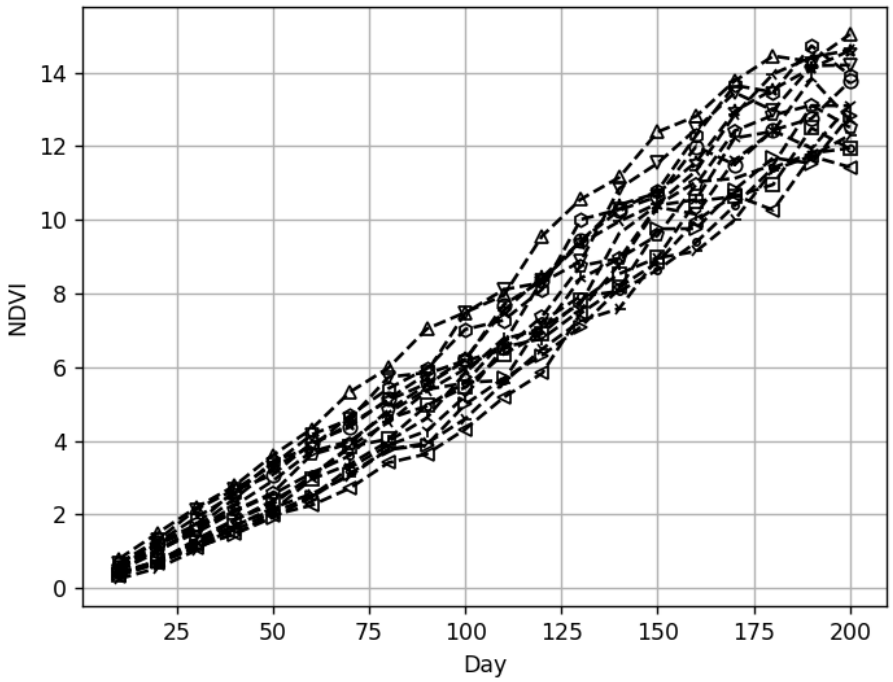


Рис. 5. Динаміка кумулятивних значень вегетаційного індексу

Апроксимація першої траєкторії за допомогою моделі Моно показала високу якість наближення з максимальною відносною похибкою 9,2% та середньою відносною похибкою 5,0%.

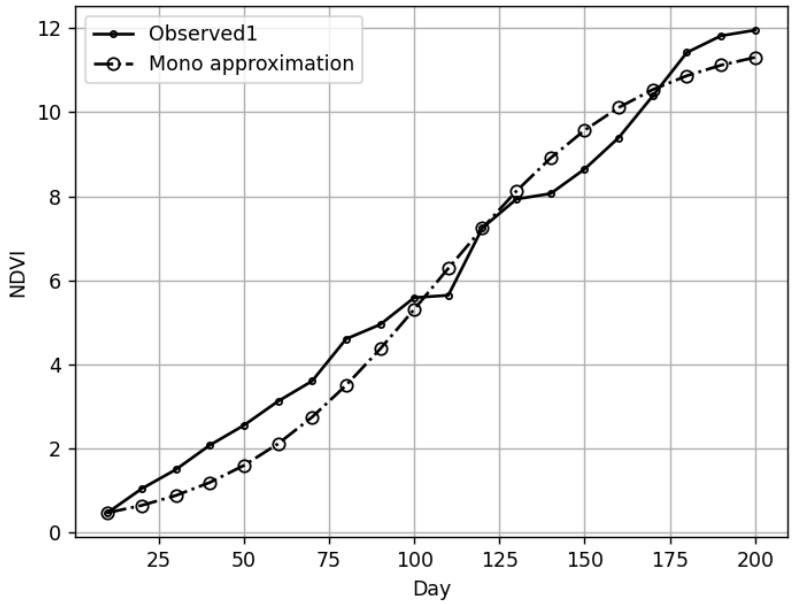


Рис.6. Апроксимація кумулятивних значень за допомогою моделі Моно

Ефективність адаптивного прогнозування була протестована на траєкторіях різної інтенсивності росту. Четверта траєкторія (нормальний ріст) показала максимальну похибку прогнозу 6,6% та середню 1,5%, а сьома траєкторія (знижена інтенсивність) – 5,2% та 1,4% відповідно. Графічна візуалізація результатів адаптивного прогнозування наведена на рис.7.

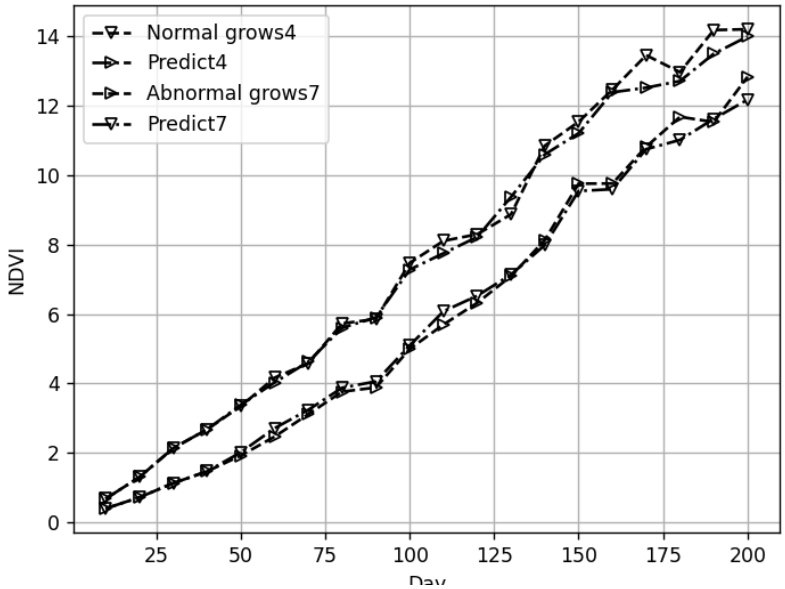


Рис. 7. Адаптивний прогноз для траєкторій різної інтенсивності

Для прогнозування урожайності пшениці були розроблені лінійна та нелінійна моделі на основі максимальних кумулятивних значень NDVI:

$$Y(f,Z) = L(\max_{d}(X_{NDVI}(f,Z,d)) \quad (5)$$

де: Y – урожайність; L – лінійний одновимірний оператор; X_{NDVI} – кумулятивне значення вегетаційного індексу $NDVI$; f – ідентифікатор поля; Z – сезон вирощування урожаю; d - день року.

Також побудуємо нелінійну залежність типу випадкового лісу

$$Y(f,Z) = RF(\max_{d}(X_{NDVI}(f,Z,d)) \quad (6)$$

де: RF – оператор побудови залежності типу випадкового лісу.

Для прогнозування урожайності пшениці були розроблені лінійна та нелінійна моделі на основі максимальних кумулятивних значень NDVI згідно з формулами (5) та (6). Результати застосування лінійної моделі, що використовувала 50% даних для навчання, демонструються на рис. 8.

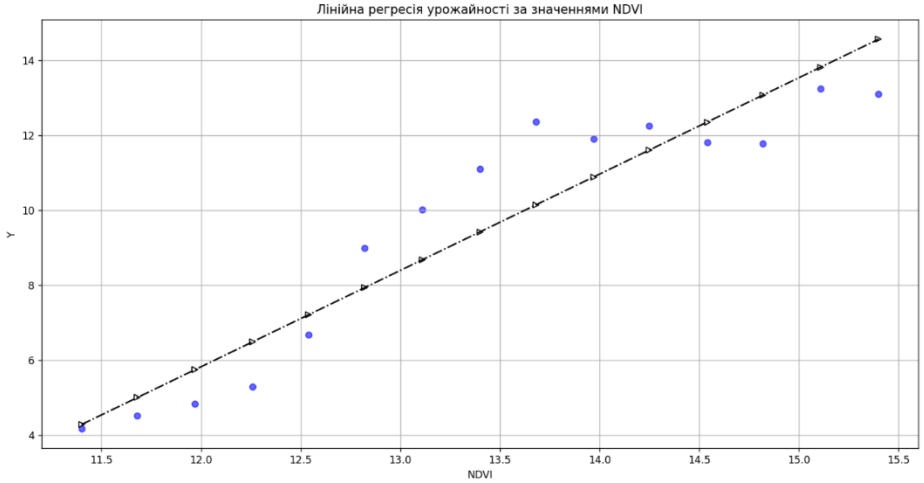


Рис. 8. Лінійна модель урожайності пшениці

Модель випадкового лісу з використанням 100 випадкових дерев показала дещо іншу структуру залежності, що відображено на рис. 9.

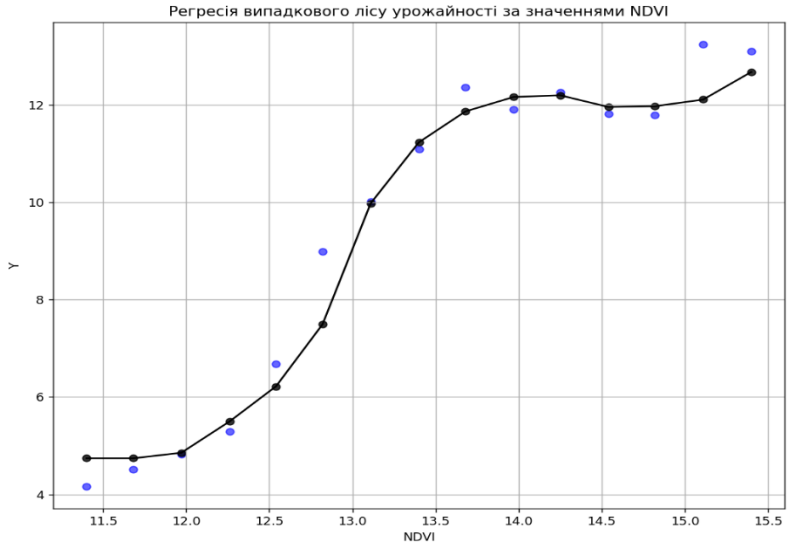


Рис. 9. Модель випадкового лісу урожайності пшениці

Результати моделювання урожайності пшениці: лінійна модель: максимальна похибка 8.9%, середня 4.8%; випадковий ліс: максимальна похибка 8.6%, середня 6.2%.

Дворівневий підхід передбачає окреме моделювання динаміки факторів урожайності (NDVI та MTCI) з наступним прогнозуванням урожайності на основі модельованих значень індексів. Для експериментальної перевірки використовувався набір з 21 траєкторії динаміки вегетаційних індексів рису в нормальних та стресових умовах. Часові профілі спостережених траєкторій NDVI характеризуються значною варіабельністю, що показано на рис. 10.

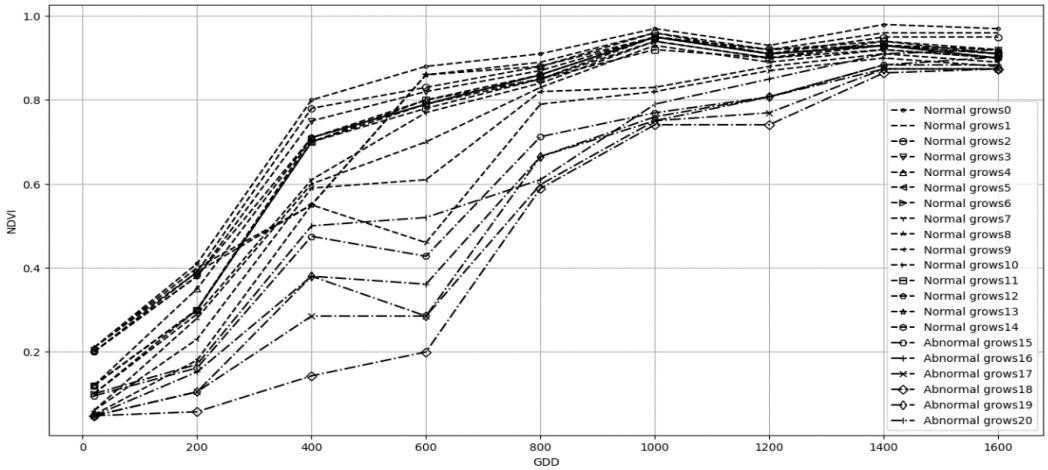


Рис. 10. Часові профілі NDVI

Аналогічно, траєкторії MTCI демонструють складну динаміку з різними патернами розвитку, що відображено на рис. 11.

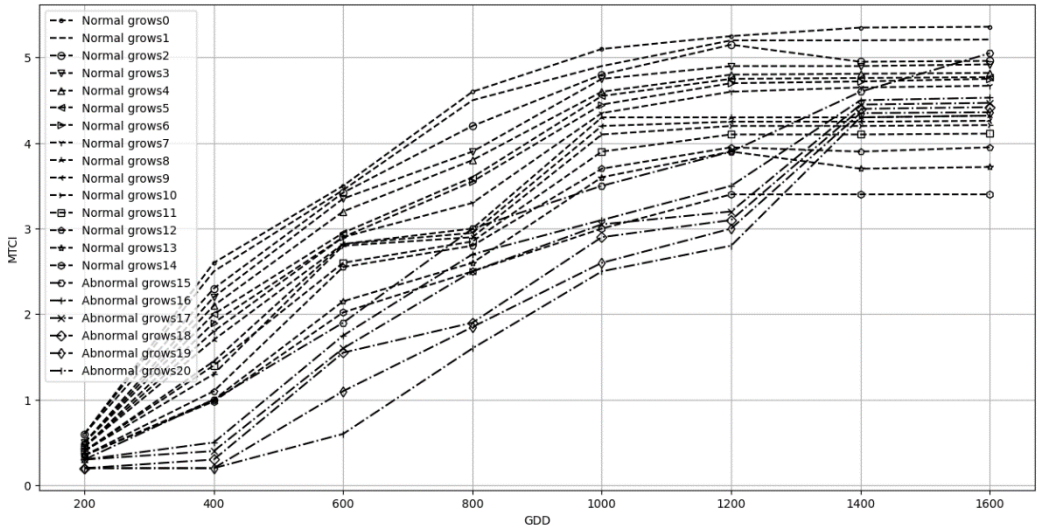


Рис. 11. Часові профілі MTCI

На базі цих даних ми можемо побудувати адаптивну дворівневу модель урожайності. Але попередньо нам потрібно дослідити її адекватність. Тому розбиваємо набори спостережених значень OV на навчальні Tr, тестувальні Ts та контрольні Cn підмножини

$$OV = Tr \cup Ts \cup Cn \tag{7}$$

На основі множини Tr будуємо адаптивні моделі вегетативних індексів NDVI, MTCI згідно поданих співвідношень (1)-(18). Далі будуємо лінійні регресійні моделі урожайності

$$Y_L = C_1 X_{NDVI} + C_2 X_{MTCI} + \mathrm{C_0}+\mathrm{C} \tag{8}$$

або в двофакторній моделі випадкового лісу

$$Y_{RF} = RF(X_{NDVI}, X_{MTCI}) . \tag{9}$$

які ідентифікуємо на основі множини тестових спостережень Ts. Далі контролюємо якість побудованих моделей на основи множини контрольних точок, які не брали участі у побудові та навчанні моделей. За результатами аналізу ефективності моделей на контрольних точках будуємо рекомендації щодо їх раціонального використання.

Результати моделювання динаміки NDVI показали достатньо високу точність для обох типів умов вирощування, що демонструється на рис. 12.

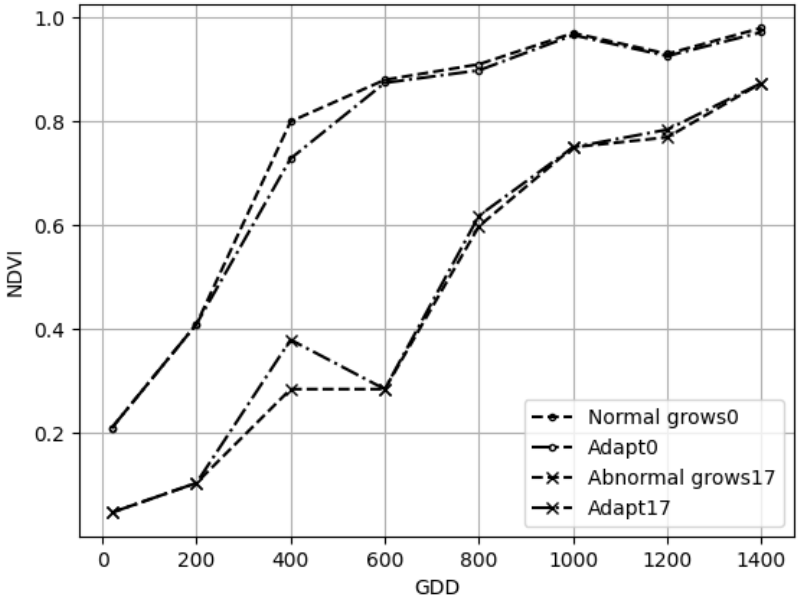


Рис. 12. Адаптивні моделі Моно для NDVI в нормальних та стресових умовах

Моделювання динаміки MTCI продемонструвало ще вищу точність, особливо для нормальних умов вирощування, як показано на рис. 13.

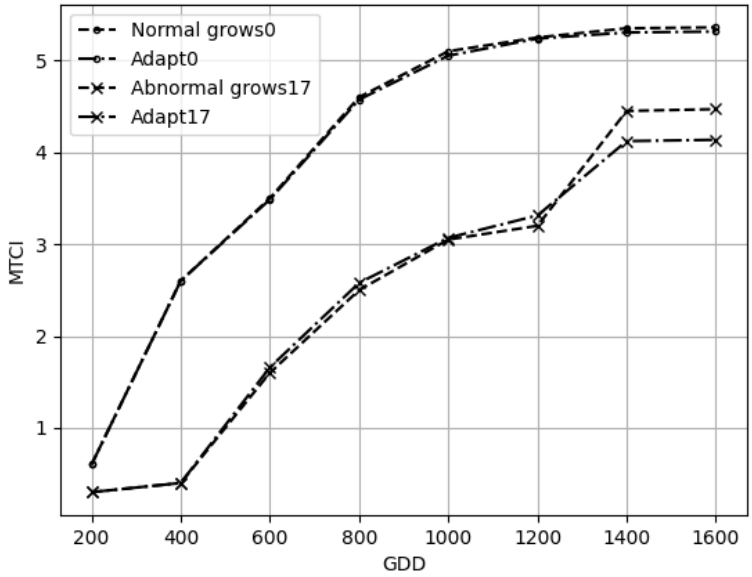


Рис.13. Адаптивні моделі Моно для MTCI в нормальних та стресових умовах

Узагальнені результати моделювання динаміки вегетаційних індексів демонструють високу ефективність запропонованого двокомпонентного адаптивного підходу. Для NDVI модель показала стабільні результати: у нормальних умовах максимальна похибка 7,3% при середній 1,3%, в стресових умовах – 10,9% та 1,8% відповідно. MTCI продемонстрував ще вищу точність з мінімальними похибками 0,9%/0,4% у нормальних умовах та 7,5%/2,6% у стресових. Отримані результати підтверджують робастність підходу та його придатність для практичного застосування в системах точного землеробства. Модель лінійної регресії для прогнозування урожайності рису показала помірну точність з чітко вираженою гіперплощиною залежності, що візуалізовано на рис. 14.

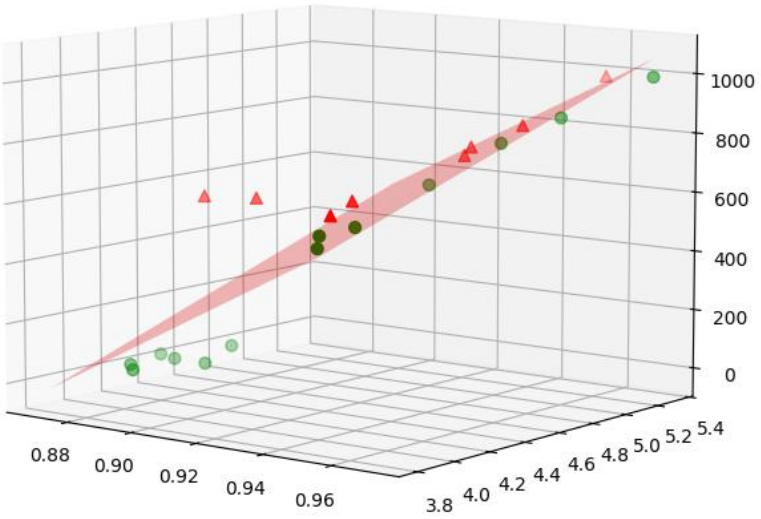


Рис. 14. 3D графік лінійної регресії урожайності

Модель випадкового лісу продемонструвала здатність до кращого відображення нелінійних залежностей між вегетаційними індексами та урожайністю, що показано на рис. 15.

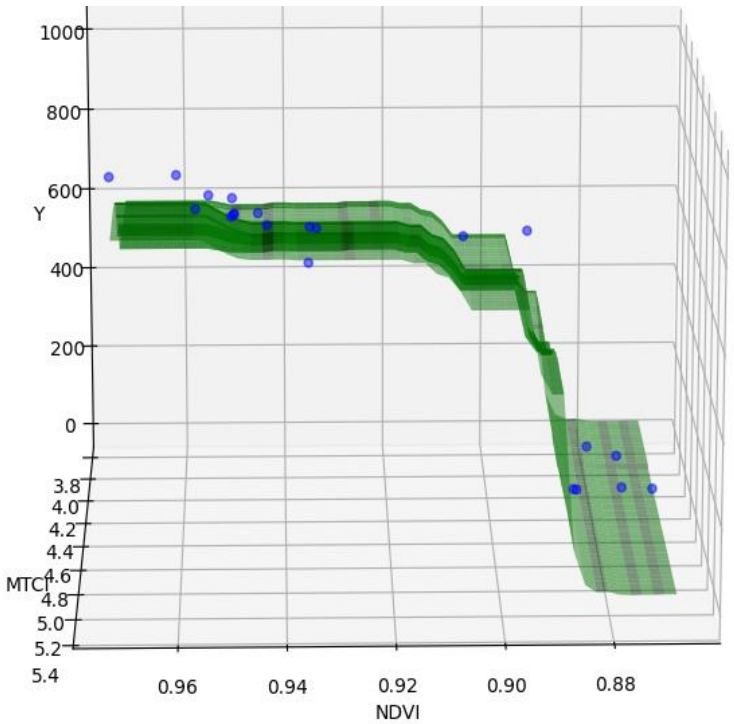


Рис. 15. 3D графік моделі випадкового лісу

Результати двофакторного моделювання урожайності рису підтверджують переваги використання нелінійних методів машинного навчання для прогнозування сільськогосподарської продуктивності. Порівняльний аналіз показав, що модель випадкового лісу демонструє суттєво вищу точність порівняно з лінійною регресією: максимальна похибка знижується з 17,7% до 14,1%, а середня похибка зменшується майже вдвічі з 8,3% до 5,3%. Це свідчить про наявність складних нелінійних взаємозв'язків між вегетаційними індексами NDVI та MTCI та кінцевою урожайністю, які краще відображаються ансамблевими методами. Отримана середня похибка 5,3% для моделі випадкового лісу є цілком прийнятною для практичного застосування в агрономічних системах підтримки прийняття рішень.

Висновки. Вперше запропоновано комплексний підхід до прогнозування урожайності сільськогосподарських культур, який базується на інноваційній двокомпонентній адаптивній моделі динаміки вегетаційних індексів. Ключовим науковим досягненням є успішне поєднання дискретних адаптивних методів з неперервним моделюванням на основі системи диференціальних рівнянь Моно, що забезпечило високу точність відтворення біологічних процесів росту рослин. Розроблена математична модель виявлення ущільнень ґрунту демонструє суттєвий прорив у підвищенні точності діагностики проблемних ділянок поля, забезпечуючи покращення результатів у півтора рази порівняно з традиційними лінійними методами. Це досягнення має важливе практичне значення для оптимізації агротехнічних заходів та підвищення ефективності використання сільськогосподарської техніки. Методологічний внесок роботи полягає у створенні дворівневої системи моделювання, де вегетаційні індекси NDVI та MTCI виступають не як прості предиктори, а як модельовані змінні з власною динамікою. Такий підхід забезпечив високу точність прогнозування з середньою похибкою 5,3% для ансамблевих методів, що є цілком прийнятним для практичного використання в агрономічних системах підтримки прийняття рішень. Технологічною основою реалізації розробленого підходу стала архітектура геоінформаційної системи на базі PostGIS, яка забезпечує ефективну обробку великих масивів просторово-часових даних від БПЛА та LiDAR систем. Експериментальна валідація на реальних агрономічних даних пшениці та рису в різних кліматичних умовах підтверджує універсальність та практичну застосовність запропонованих методів для розвитку систем точного землеробства.

Список літератури

1. Basso B., Liu L. Seasonal crop yield forecast: Methods, applications, and accuracies. *Advances in Agronomy*. 2019. T.154. C. 201-255.
2. Xue J., Su B. Significant remote sensing vegetation indices: A review of developments and applications. *Journal of Sensors*. 2017. 1353691.
3. Lobell D.B., Burke M.B. On the use of statistical models to predict crop yield responses to climate change. *Agricultural and Forest Meteorology*. 2010. T.150 (11). C.1443-1452.
4. Reynolds M., Dreccer F., Trethowan R. Drought-adaptive traits derived from wheat wild relatives and landraces. *Journal of Experimental Botany*. 2007. T.58(2). C. 177-186.
5. Pasichnyk R. M., Babala, L. V., & Machuliak, M. V. A Method for Improving the Quality of Image Annotation in Semantic Monitoring Gis of Business Processes. Informatics and Mathematical Methods in Simulation. 2024. T.14(3).
6. Pasichnyk R., Babala L., Machulyak M. Vegetation Indices Dynamics Model in GIS Based on an Adaptive Predictive Method and the Mono System. *15th International Conference on Advanced Computer Information Technologies (IEEE)* . 2025. P. 186-191).
7. Jones J.W. The DSSAT cropping system model. *European Journal of Agronomy*. 2003. V.18 (3-4). P. 235-265.
8. 8.Zhang W., Liu L., Wang T. Deep learning approaches for crop yield prediction using satellite imagery. *Remote Sensing*. 2024. V.16 (4). P. 875.
9. Monod J. Recherches sur la croissance des cultures bactériennes. Paris: Hermann & Cie, 1942

TWO-COMPONENT ADAPTIVE MODEL OF VEGETATION INDICES DYNAMICS FOR AGRICULTURAL CROP YIELD PREDICTION

M.V. Machulyak

Western Ukrainian National University
11, Lvivska Str., Ternopil, 46009, Ukraine
Email: Mvmach9@gmail.com

The paper proposes a two-component adaptive model of vegetation indices dynamics, which, unlike existing ones, contains a discrete adaptive component based on a combination of current and historical data and a continuous approximation component based on the Monod differential equations system. A method for identifying a two-level adaptive yield model has been developed that uses explanatory variables in the form of vegetation indices models NDVI and MTCI instead of their direct values. The proposed geographic information system architecture based on PostGIS provides efficient integration of spatial data from UAV and LiDAR systems. A mathematical model for detecting soil compaction areas based on Monod equations showed a 1.5-fold accuracy improvement compared to linear methods. Experimental validation on wheat and rice data confirmed the approach effectiveness with an average yield prediction error of 5.3% for ensemble methods. The obtained results demonstrate the promising application of the proposed approach for precision agriculture systems development and agricultural decision-making support.

Keywords: vegetation indices, Monod model, yield prediction, adaptive modeling, remote sensing, geographic information systems, precision agriculture

DOI 10.15276/imms.v15.no4.608
УДК 519.6, 539.3

Informatics and Mathematical Methods in Simulation
Vol.15 (2025), No. 4, pp. 608-613

МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ МЕТОДОМ РЯДІВ ТРИВИМІРНОЇ КРАЙОВОЇ ЗАДАЧІ ДЛЯ КОСОСИМЕТРИЧНОГО ЗГИНАННЯ ШАРУ З КРУГОВИМ ОТВОРОМ

Б.Є. Панченко[1], Ю.Д. Ковальов[2], Л.М. Тимошенко[3], Г.О. Фесенко[4], М.В. Северин[5]

[1,2,5]Державний університет інтелектуальних технологій та зв'язку
1, Кузнечна вул., Одеса, 65023, Україна
[3]Національний університет «Одеська політехніка»
1, Шевченка пр., Одеса, 65044, Україна
[4]Одеський національний університет ім. І. І.Мечникова
2, Змієнка Всеволода вул., Одеса, 65082, Україна
Emails: pr-bob@ukr.net[1], kovalev@ukr.net[2], fesenko@onu.edu.ua[4], n_severin@ukr.net[5]

Розглянуто мішану задачу теорії пружності для нескінченного шару з отвором-порожниною вздовж висоти шару, коли на плоских гранях задано умови гладкого контакту, а по поверхні отвору діє навантаження. Невідомі переміщення, що виникають у шарі розкладаються у ряди Фур′є і методом розділення змінних система рівнянь рівноваги зводиться до системи диференціальних рівнянь відносно невідомих гармонік. Розв′язки розшукуються у вигляді комбінацій метагармонічних функції та їхніх похідних. Для визначення цих функцій отримана система лінійних алгебраїчних рівнянь. В результаті отримане окружне напруження, що виникає у шарі, було проаналізовано в залежності від матеріалу шару, вигляду отвору та співвідношень між радіусом отвору та висотою шару.

Ключові слова: тривимірна крайова задача, пружний шар, наскрізний отвір, метод рядів, чисельний експеримент.

Вступ. Математичне моделювання статичних тривимірних крайових задач у теорії пружності є важливим напрямом механіки деформівного твердого тіла, що має широке прикладне значення в інженерії та матеріалознавстві. Зокрема, задачі, пов’язані з тілами, які мають отвори, є актуальними через наявність концентрації напружень у таких зонах, що істотно впливає на міцність та надійність конструкцій.

Одним з ефективних аналітичних підходів до розв’язання подібних задач є метод рядів. Він дозволяє звести систему диференціальних рівнянь рівноваги до сукупності алгебраїчних співвідношень шляхом подання шуканих функцій напружень та переміщень у вигляді збіжних рядів за спеціальними функціями – сферичними чи циліндричними гармоніками, які задовольняють основним рівнянням теорії пружності та крайовим умовам.

Класичне дослідження Грінченка і Улітка [1] заслуговує на особливу увагу, оскільки в ньому було отримано точний аналітичний розв’язок задачі про розподіл напружень поблизу кругового отвору у пружному шарі. Запропонований підхід дозволив врахувати геометричні особливості задачі та забезпечити високу точність розрахунків у критичних зонах концентрації напружень, що має важливе значення для оцінки міцності та довговічності конструкцій із отворами.

У монографії [2] Поповим Г.Я. розглянуто задачі теорії пружності, пов’язані з концентрацією напружень навколо неоднорідностей, таких як тріщини та тонкі включення. Автором запропоновано методи аналітичного та наближеного розв’язання задач для різних областей та типів навантажень. Значну увагу приділено побудові асимптотичних розв’язків і аналізу напружено-деформованого стану в околі неоднорідностей. Робота є важливою для розробки та оцінки міцності елементів конструкцій з урахуванням ефектів концентрації напружень.

В роботі [3] Фесенко та Вайсфельд досліджують динамічну задачу для нескінченного пружного шару з циліндричною порожниною. Застосовано метод інтегральних перетворень, що дозволяє точно описати хвильові процеси в присутності геометричних неоднорідностей.

В роботі Панченка, Ковальова та співавт. [4] розглянуто кососиметричну задачу для пружного шару з наскрізним отвором, в якій використовуються крайові умови ковзного защемлення торців [5]. Авторами запропоновано аналітичний підхід на основі подання розв'язку у вигляді рядів спеціальних функцій, що дозволяє дослідити розподіл напружень та переміщень у шарі.

Стаття [6] розширює підхід до моделювання тривимірних задач для пружного шару з неоднорідностями. Розглядаються різні крайові умови на торцях, що дозволяє моделювати широкий клас практичних ситуацій. В роботі приділено увагу побудові розв'язків із застосуванням розвинень за ортогональними функціями.

Істотний внесок у моделювання шарів із наскрізними отворами зроблено у працях Фільштинського та Ковальова [7], де проаналізовано напружений стан п'єзокерамічного шару зі складною геометрією. Розв'язок побудовано з урахуванням п'єзоелектричних ефектів, що розширює область застосування моделі для функціональних матеріалів.

В роботі [8] запропоновано фундаментальні розв'язки для кососиметричного випадку в шарі типу R^3 з урахуванням змішаних крайових умов. Представлені розв'язки можуть бути використані як базисні функції для побудови наближених розв'язків задач із більш складною геометрією та навантаженням.

Отже, наведені результати засвідчують активний розвиток методів математичного моделювання для тривимірних крайових задач, зокрема із застосуванням рядів і спеціальних функцій, що відкриває можливості для точного опису поведінки тіл із отворами та неоднорідностями.

В межах цієї роботи розглядається побудова аналітичного розв'язку методом рядів [1] крайової задачі про кососиметричний вигін пружного шару з круговим отвором та ковзним защемленням торців [5] для опису його напружено-деформованого стану (НДС). Досліджується вплив матеріалу та геометричних параметрів шару, а також радіусу отвору на розподіл напружень. Проведено порівняння отриманих тут результатів з результатами, отриманих методом сингулярних інтегральних рівнянь [4, 6].

Надані результати можуть бути використані для підвищення точності розрахунків конструктивних елементів з отворами та оптимізації їх форми з урахуванням напруженого стану.

Постановка задачі. Розглянемо пружний шар $-h \le x_3 \le h$, $-\infty < x_1, x_2 < \infty$, послаблений наскрізним отвором вздовж осі Ox_3, на поверхні якого діє поверхневе навантаження (N, T, Z), а на плоских гранях шару $x_3 = \pm h$ задано умови гладкого контакту. Переміщення $u_n(x_1, x_2, x_3)$, $n = 1, 2, 3$, що виникають у шарі задовольняють рівнянням рівноваги

$$\Delta \mathbf{u} + \sigma \, grad\, \theta = 0, \tag{1}$$

де Δ – оператор Лапласа, $\mathbf{u} = \left(u_1(x_1, x_2, x_3), u_2(x_1, x_2, x_3), u_3(x_1, x_2, x_3)\right)$ – вектор переміщень, $\theta = \dfrac{\partial u_1}{\partial x_1} + \dfrac{\partial u_2}{\partial x_2} + \dfrac{\partial u_3}{\partial x_3}$ – об'ємне розширення, $\sigma = (1 - 2\nu)^{-1}$, ν – коефіцієнт Пуассона. Граничним умовам на плоских торцях шару мають вигляд:

$$u_3(x_1, x_2, \pm h) = 0, \sigma_{13}(x_1, x_2, \pm h) = \sigma_{23}(x_1, x_2, \pm h) = 0 \tag{2}$$

Запишемо компоненти вектору переміщення у вигляді рядів Фур'є

$$u_i = \sum_{k=0}^{\infty} u_{ik}(x_1, x_2)\sin\gamma_k x_3,\ i=1,2,\ u_3 = \sum_{k=0}^{\infty} u_{3k}(x_1, x_2)\cos\gamma_k x_3,\ \gamma_k = \frac{2k+1}{2h}\pi, \tag{3}$$

які автоматично задовольняють умовам (2) на торцях шару. Після підстановки подань (3) у рівняння рівноваги (1) та розділення змінних отримаємо систему відносно компонент $u_{ik}(x_1, x_2)$, $i=1,2,3$

$$\kappa_k u_{ik} + \sigma\partial_i\theta_k = 0,\ i=1,2,$$

$$\kappa_k u_{3k} + \sigma\gamma_k\theta_k = 0, \tag{3}$$

де

$$\kappa_k = \nabla^2 - \gamma_k^2,\ \nabla^2 = \partial_1^2 + \partial_2^2,\ \theta_k = \partial_1 u_{1k} + \partial_2 u_{2k} - \gamma_k u_{3k},\ \partial_i = \partial/\partial x_i,\ i=1,2.$$

Безпосередньо з системи (3) знайдено, що

$$\kappa_k\theta_k = 0 \Rightarrow \left(\nabla^2 - \gamma_k^2\right)\theta_k = 0, \tag{4}$$

що означає метагармонічність функції θ_k. Виходячи з цього, введемо функцію ψ_k співвідношенням:

$$\theta_k = \kappa_k\psi_k \tag{5}$$

Із співвідношення (4) маємо, що $\kappa_k^2\psi_k = 0$. З урахуванням зв′язку (5) між функціями ψ_k та θ_k інтегрування системи (3) дає можливість отримати

$$u_{1k} = -\sigma\partial_1\psi_k + \varphi_{1k},\ u_{2k} = -\sigma\partial_2\psi_k + \varphi_{2k},\ u_{3k} = -\gamma_k\sigma\psi_k + \varphi_{3k}, \tag{6}$$

де $\kappa_k\varphi_{ik} = 0,\ i=1,2$.

Після вимоги фактичного виконання співвідношення (5), отримаємо, що

$$\varphi_{1k} = \sigma\partial_2\varphi_k,\ \varphi_{2k} = -\sigma\partial_1\varphi_k,\ \varphi_{3k} = -\frac{1+\sigma}{\gamma_k}\kappa_k\psi_k, \tag{7}$$

де φ_k – довільний розв′язок рівняння $\kappa_k\varphi_{ik} = 0,\ i=1,2$.

Формули (3), (6), (7) дають вирази пружних переміщень у шарі через функції φ_k, ψ_k.

В силу (6), (7) маємо

$$u_{1k} - iu_{2k} = 2\sigma\frac{\partial}{\partial z}(i\varphi_k - \psi_k),\ \ u_{3k} = -\left(\frac{1+\sigma}{\gamma_k}\kappa_k + \sigma\gamma_k\right)\psi_k, \tag{8}$$

де $\dfrac{\partial}{\partial z} = \dfrac{1}{2}(\partial_1 - i\partial_2),\ z = x_1 + ix_2$.

Граничні умови на контурі L отвору запишемо у комплексній формі

$$\begin{aligned}&(\sigma_{11}+\sigma_{22}) - e^{2i\psi}(\sigma_{22}-\sigma_{11}+2i\sigma_{12}) = 2(N-iT)\\ &\mathrm{Re}\left[e^{-i\psi}(\sigma_{13}+i\sigma_{23})\right] = Z\end{aligned} \tag{9}$$

де ψ – кут між зовнішньою нормаллю до контуру порожнини-отвору та віссю $0x_1$.

Використовуючи закон Гуку та формули (6), представимо умови (9) у формі

$$\begin{aligned}&2\sigma e^{2i\psi}\left\{\frac{\partial^2}{\partial z^2}(i\varphi_k - \psi_k)\right\} - \frac{1}{2}\theta_k - \frac{1}{2}\sigma\gamma_k^2\psi_k = \frac{1}{2\mu}(N_k - iT_k),\\ &-\mathrm{Re}\left\{e^{i\psi}\left[\frac{\partial}{\partial z}\left(\sigma\gamma_k\psi_k + \frac{1+\sigma}{\gamma_k}\theta_k\right) - \sigma\gamma_k\frac{\partial}{\partial z}(i\varphi_k - \psi_k)\right]\right\} = \frac{1}{2\mu}Z_k,\end{aligned} \tag{10}$$

де μ – модуль зсуву.

Функцій φ_k и ψ_k розшукуємо у вигляді:

$$\varphi_k = 0,\ \psi_k = A_k K_0(\gamma_k r) + B_k r\, K_1(\gamma_k r),$$

де A_k, B_k – невідомі константи, $K_n(\gamma_k r)$ – функції Макдональда порядку $n = 0,1$, $r = |\zeta - z|,\ \zeta = \xi + i\eta \in L,\ z = x_1 + ix_2$.

Підставивши вирази у граничні умови, та розглянувши випадок $T_k = 0,\ Z_k = 0$ на $r = 1$, отримаємо систему лінійних алгебраїчних рівнянь (СЛАР)

$$a_{11}A_k + a_{12}B_k = \frac{1}{2\mu} N_k$$

$$a_{21}A_k + a_{22}B_k = 0$$

де

$$a_{11} = -\sigma\gamma_k K_1(\gamma_k r) - \sigma\gamma_k^2 K_0(\gamma_k r),\ a_{12} = \gamma_k K_0(\gamma_k r) - \sigma\gamma_k^2 K_1(\gamma_k r),$$

$$a_{21} = \sigma\gamma_k K_1(\gamma_k r),\ a_{22} = \sigma\gamma_k K_0(\gamma_k r) - (1+\sigma) K_1(\gamma_k r).$$

Звідки знайдено, що

$$A_k = N_k \frac{a_{22}}{d},\ B_k = -N_k \frac{a_{21}}{d},\ d = a_{11}a_{22} - a_{12}a_{21},\ N_k = \frac{N_k}{2\mu},$$

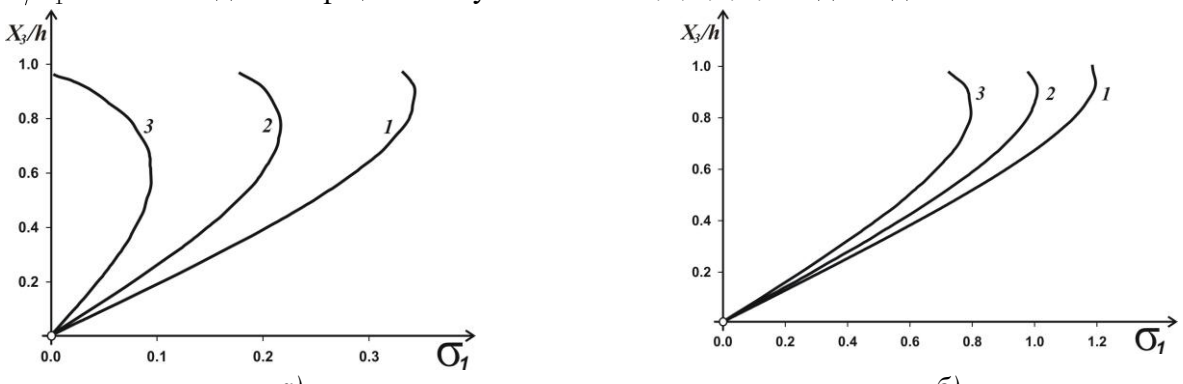
$$\sigma_\theta^k = N_k \left(\frac{a_{22}d_1}{d} - \frac{a_{21}d_2}{d} - 1 \right),\ d_1 = -\sigma\gamma_k^2 K_0(\gamma_k r),\ d_2 = 2\gamma_k K_0(\gamma_k r) - \sigma\gamma_k^2 K_1(\gamma_k r),$$

В результаті маємо остаточний вираз для окружного напруження, що виникає у шарі

$$\sigma_{\theta\theta} = \sum_{k=1}^{\infty} \sigma_\theta^k \sin(\gamma_k x_3).$$

Результати чисельного дослідження. Розглянемо випадок, коли на поверхні отвору діє навантаження $N = -Px_3,\ T = Z = 0,\ P = const$.

На рис. 1 наведені епюри розподілу відносного окружного напруження $\sigma_1 = \sigma_{\theta\theta}/P$ вздовж координати в точці $\varphi = 0$ кругового отвору $(R_1 = R_2 = 1)$ за товщиною. Криві 1, 2, 3 рис. 1а побудовані для $h/R_1 = 1$ та коефіцієнтів Пуассона $\nu = 0{,}2;\ 0{,}3;\ 0{,}4$ відповідно. Ці результати порівнювалися з результатами робот [3,5]. Збіг показує достовірність використаного тут методу. Криві 1;2;3 рис. 1б наведено для $h/R_1 = 2$ також для коефіцієнтів Пуассона $\nu = 0{,}2;\ 0{,}3;\ 0{,}4$ відповідно.

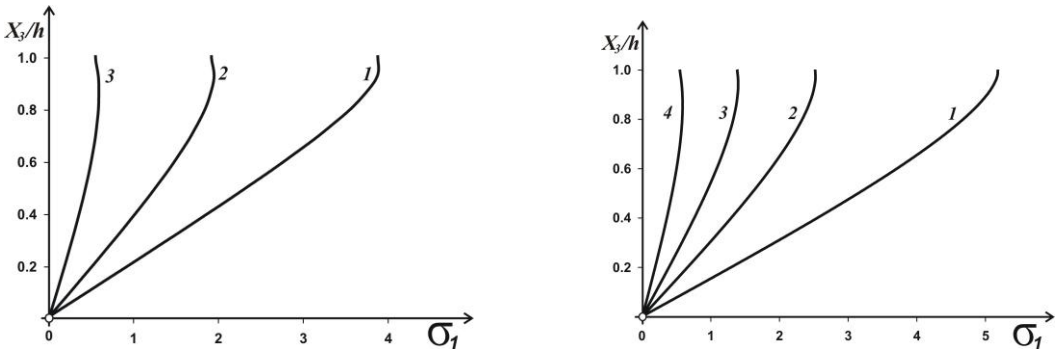


а) *б)*

Рис. 1. Епюри розподілу відносного окружного напруження вздовж координати в точці $\varphi = 0$ кругового отвору за товщиною

На рис. 2а наведені епюри розподілу відносного окружного напруження $\sigma_1 = \sigma_{\theta\theta}/P$ вздовж координати в точці $\varphi = 0$ за товщиною. Криві 1, 2, 3 побудовані для кругового отвору $(R_1 = R_2 = 1)$ та коефіцієнту Пуассона $\nu = 0{,}3$, де $h/R_1 = 1{,}5;\ 3;\ 5$ відповідно.

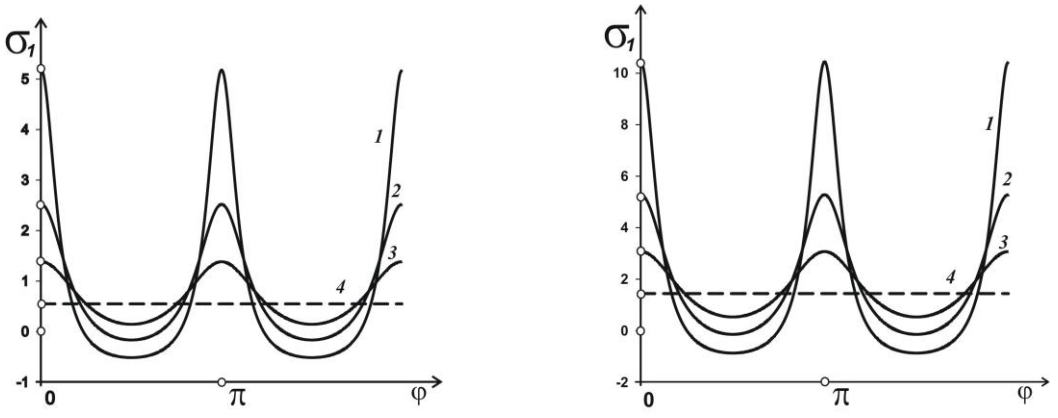
На рис. 2б наведені епюри розподілу відносного окружного напруження $\sigma_1 = \sigma_{\theta\theta}/P$ вздовж координати в точці $\varphi = 0$ за товщиною для еліптичного отвору у випадку $h/R_1 = 1$ $(R_1 = 1)$. Криві 1;2;3;4 наведено для $R_2 = 0{,}3; 0{,}5; 0{,}7; 1$ відповідно



а) *б)*

Рис. 2. Епюри розподілу відносного окружного напруження вздовж координати в точці $\varphi = 0$ за товщиною для еліптичного отвору

На рис. 3 наведені епюри розподілу відносного окружного напруження $\sigma_1 = \sigma_{\theta\theta}/P$ по контуру еліптичного отвору $0 \le \varphi \le 2\pi$ а площині $x_3/h = 1$ у випадку $h/R_1 = 1{,}5;\ 2{,}5$ $(R_1 = 1)$ на рис. 3а та 3б відповідно. Для отримання нових чисельні результатів для отвору з еліптичним контуром використано математичну модель, запропоновану в [3,5]. Криві 1; 2; 3; 4 наведені відповідно до $R_2 = 0{,}3; 0{,}5; 0{,}7; 1$. За формую та значеннями напружень отримано очікувані результати. Як і в [3,5] для отвору кругової форми криві 4 вироджуються в прямі, що підтверджується і використанням алгоритму розв'язку задачі в рядах. Тут чисельні результати отримані зі значенням коефіцієнта Пуассона $\nu = 0{,}3$.



а) *б)*

Рис. 3. Епюри розподілу відносного окружного напруження по контуру еліптичного отвору

Висновки. Запропонована математична модель розв'язання тривимірних статичних крайових задач для шару з наскрізним отвором та ковзному защемленні торців, є адекватною та ефективною. Алгоритм добре масштабується та надає можливість ефективного керування ресурсами.

Для поставленої задачі за результатами чисельного дослідження можна зробити такі висновки.

1. Зі збільшенням товщини шару відбувається зростання відносного напруження.

2. У разі кругового отвору спостерігається зменшення максимуму відносного окружного напруження від торців у глибину шару.
3. Зі зменшенням R_2 спостерігається зростання відносного окружного напруження.

Список літератури

1. Grinchenko V. T., Ulitko A. F. An exact solution of the problem of stress distribution close to a circular hole in an elastic layer. *Sov. Appl. Mech.* 1968. Vol.4. No.10. P. 31-37.
2. Попов Г. Я. Концентрация упругих напряжений возле штампов, разрезов, тонких включений и подкреплений. М.: Наука, 1982. 344 с.
3. Fesenko A., Vaysfel'd N. The dynamical problem for the infinite elastic layer with a cylindrical cavity. *Procedia Structural Integrity.* 2021. Vol. 33. P. 509-527.
4. Панченко Б.Є., Ковальов Ю.Д., Каліна Т.О., Сайко І.М., Буката Л.М. Математичне моделювання в статичних тривимірних крайових задачах – кососиметрична задача для шару, послабленого наскрізним отвором при ковзному защемленні торців. *Кібернетика та системний аналіз*. 2024. Т.60. №1. С. 182-195.
5. Шевченко В.П., Алтухов Е.В., Фоменко М.В., Деформация трехслойных пластин со скользящей заделкой торцов и несовершенным контактом слоев. *Доповіді Національної академії наук України.* 2012. № 8. С. 61-66
6. Панченко Б.Є., Ковальов Ю.Д., Буката Л.М. Северин М.В. Математичне моделювання деяких тривимірних крайових задач для шару з неоднорідностями та різними крайовими умовами на торцях. *Colloquium-journal. Computer science*. 2024. №13 (206). С. 19-31.
7. Фільштинський Л.А., Ковальов Ю.Д. Моделювання напруженого стану п'єзокерамічного шару, ослабленого наскрізними тунельними отворами. *Вісник Херсонського державного технічного університету.* 2000. №2 (8). С. 216-219.
8. Фільштинський Л.А., Шрамко Л.В. Фундаментальні рішення для п'єзокерамічного шару R3 (кососиметричний випадок, змішані граничні умови). *Теорет. та прикладна механіка*. 2003. Вип. 38. С. 53-57.

MATHEMATICAL MODELING BY SERIES METHOD OF A THREE-DIMENSIONAL BOUNDARY-VALUE PROBLEM FOR SKEW-SYMMETRIC BENDING OF A LAYER WITH A CIRCULAR HOLE

B.E. Panchenko[1]., Yu.D. Kovalev[2], L.M. Timoshenko[3], G.O. Fesenko[4], M.V. Severyn[5]

[1,2,5]State University of Intellectual Technologies and Communications
1, Kuznechna Str., Odesa, 65023, Ukraine
[3]National Odesa Polytechnic University
1, Shevchenko Ave., Odesa, 65044, Ukraine
[4]Odesa National University named after I. I. Mechnikov
2, Zmienko Vsevolod Str., Odesa, 65082, Ukraine
Emails: pr-bob@ukr.net[1], kovalev@ukr.net[2], fesenko@onu.edu.ua[4], n_severin@ukr.net[5]

A mixed problem of the theory of elasticity for an infinite layer with a hole-cavity along the height of the layer is considered, when the conditions of smooth contact are given on the flat faces, and a load acts on the surface of the hole. The unknown displacements arising in the layer are expanded into Fourier series and by the method of separation of variables the system of equilibrium equations is reduced to a system of differential equations with respect to unknown harmonics. Solutions are sought in the form of combinations of metaharmonic functions and their derivatives. To determine these functions, a system of linear algebraic equations is obtained. The resulting circumferential stress arising in the layer was analyzed depending on the layer material, the shape of the hole, and the relationship between the hole radius and the layer height.

Keywords: three-dimensional boundary value problem, elastic layer, through hole, series method, numerical experiment.

DOI 10.15276/imms.v15.no4.614
УДК 004.056.5

Informatics and Mathematical Methods in Simulation
Vol.15 (2025), No. 4, pp. 614-624

ПІДВИЩЕННЯ СТІЙКОСТІ СУЧАСНИХ БЛОКОВИХ ШИФРІВ ЗА ДОПОМОГОЮ ВИСОКОЯКІСНИХ S-БЛОКІВ

В. В. Радуш

Національний університет «Одеська політехніка»
1, Шевченка пр., Одеса, 65044, Україна
Email: radush9860@gmail.com

Забезпечення високого рівня криптографічної стійкості сучасних симетричних шифрів безпосередньо залежить від якості застосованих у них S-блоків. Традиційно аналіз S-блоків обмежується апаратом булевих функцій, однак останні дослідження показують, що в умовах атак на основі багатозначної логіки окремі стандартні S-блоки можуть демонструвати недостатній рівень криптографічної стійкості. Це створює необхідність комплексної оцінки таких криптографічних примітивів для забезпечення їх якості як у булевому, так і в багатозначному поданні. Проведена експериментальна оцінка якості S-блоків, що застосовуються в сучасних симетричних блокових шифрах, з акцентом на практичну ефективність при реальному шифруванні та оцінку стохастичних властивостей криптограми. З метою посилення криптографічних характеристик шифрів, замість оригінальних підстановок були використані авторські S-блоки, побудовані на основі недвійкових афінних перетворень, які демонструють високі криптографічні властивості при представленні як булевими функціями, так і функціями багатозначної логіки. Методика включала два взаємодоповнюючих етапи: розрахунок ключових метрик якості S-блоків – нелінійності, виконання суворого лавинного критерію (SAC), критерію незалежності бітів (BIC), ймовірності лінійної (LAP) і диференціальної (DAP) апроксимації; практичну верифікацію шляхом шифрування інформаційних масивів модифікованими шифрами та подальшого статистичного аналізу вихідних криптограм за допомогою NIST-тестів. Дослідження охопило криптоалгоритми AES, Camellia, Kalyna, SM4 і ARIA. Розрахункові показники разом зі стохастичною перевіркою показали, що S-блоки на основі четвіркових афінних перетворень забезпечують підвищену збалансованість і стійкість. У більшості випадків модифіковані версії шифрів випередили оригінали за кількістю пройдених NIST-тестів (зокрема AES, ARIA, SM4, Camellia), що свідчить про посилення випадковості та криптостійкості вихідних послідовностей. Важливо підкреслити експериментальний характер роботи: проведена порівняльна та практична перевірка S-блоків на основі афінних перетворень. Отримані дані вказують на перспективність подальшої інтеграції таких примітивів у протоколи зв'язку, захист IoT-пристроїв, хмарні сервіси та критичну інфраструктуру, де вимоги до випадковості та стійкості особливо високі. Поєднання теоретичних критеріїв якості та практичної перевірки криптограм дозволяє говорити про надійну методологічну базу для прийняття рішень щодо впровадження таких S-блоків у реальні системи.

Ключові слова: криптографічні примітиви; блокові шифри; S-блоки; афінні перетворення; багатозначна логіка; нелінійність; лавинний критерій; NIST Statistical Test Suite; криптоаналіз; випадковість криптограм.

Вступ та постановка проблеми. Цифрові технології сьогодні є невід'ємною складовою сучасного суспільства, оскільки використання комп'ютерних систем та мережі Інтернет перетворилося з факультативної можливості на обов'язкову умову ефективної діяльності у більшості сфер. Значна частина щоденних процесів – від професійної діяльності до побутових завдань – тісно інтегрована з цифровим середовищем. У результаті, переважна більшість інформації у світі існує у цифровому вигляді, що забезпечує можливість її довготривалого зберігання, масштабної обробки та оперативної передачі у глобальних масштабах.

Водночас із беззаперечними перевагами цифровізації постають і суттєві виклики, пов'язані з інформаційною безпекою. Персональні дані, що циркулюють у кіберпросторі,

можуть стати об’єктом несанкціонованого доступу, викрадення чи зловживання з боку зловмисників. Такі загрози актуалізують необхідність розроблення та впровадження ефективних механізмів захисту інформації. Саме цим завданням покликані відповідати криптографічні методи та комплексні засоби мережевої безпеки [1].

Криптографія – це наукова дисципліна, що вивчає методи та засоби перетворення інформації з метою забезпечення її конфіденційності, цілісності, автентичності та невідмовності. Основним завданням криптографії є трансформація відкритого (читабельного) тексту у зашифрований (шифротекст) таким чином, щоб без відповідних криптографічних реквізитів відновлення початкових даних було неможливим. При цьому за наявності ключа чи іншої секретної інформації передбачена можливість виконати зворотне перетворення – відновлення оригінального повідомлення.

Усі сучасні криптографічні алгоритми умовно поділяються на дві фундаментальні категорії – симетричну та асиметричну криптографію. Симетричні методи, у свою чергу, охоплюють кілька підтипів: блокові шифри, потокові шифри та криптографічні геш-функції [2].

Однак, незалежно від рівня складності криптографічного алгоритму, його основу становлять дві фундаментальні операції – підстановка та перестановка. Вони відомі ще з часів класичної криптографії, однак саме Клод Шеннон надав їм сучасного теоретичного обґрунтування, визначивши їх як ключові механізми для досягнення конфузії та дифузії у криптосистемах. Підстановка має фактично вирішальну роль майже в усіх сучасних криптографічних алгоритмах, таких як, наприклад, визнаний стандарт AES. Досягається ця операція за допомогою використання відповідних криптографічних примітивів – блоків підстановки, що називаються S-блоками. S-блок – це базовий нелінійний компонент криптографічних алгоритмів, що реалізує відображення елементів вхідного алфавіту у вихідний алфавіт. Його основною функцією є забезпечення конфузії, тобто ускладнення зв’язку між ключем і шифротекстом, що значно підвищує стійкість криптосистеми до криптоаналітичних атак [3]. Очевидно, що якість криптографічних примітивів безпосередньо визначає стійкість та ефективність усього криптоалгоритму, який їх застосовує. Зокрема, чим вищий рівень реалізованих у S-блоці конфузії та дифузії, тим вищою буде криптографічна якість і надійність алгоритму загалом.

Для оцінки властивостей S-блоків використовують наступні метрики:

1. Відстань нелінійності.
2. Критерій незалежності бітів (BIC).
3. Відповідність суворому лавинному критерію (SAC).
4. Статистична незалежність виходу S-блоку підстановки від його входу (кореляційний імунітет).
5. Ймовірність лінійної апроксимації (LAP).
6. Ймовірність диференціальної апроксимації (DAP).

Варто зазначити, що оцінка якості S-блоків у більшості випадків здійснюється виключно в межах апарату булевих функцій, ігноруючи альтернативні підходи до їх подання. Водночас існує ненульова ймовірність атак, що базуються на апараті багатозначної логіки, у межах яких криптографічні властивості обраного S-блоку можуть виявитися істотно слабшими [4].

Сьогодні відомо багато підходів до синтезу високоякісних S-блоків, зокрема авторські методи, представлені у відповідних дослідженнях. Побудовані за цими методами S-блоки демонструють високий рівень відповідності формалізованим метрикам, що використовуються для оцінювання криптографічних властивостей, таких як нелінійність, стійкість до диференціального та лінійного криптоаналізу тощо.

Разом із тим, залишається відкритим питання практичної ефективності таких S-блоків у складі конкретних криптографічних алгоритмів. Попри високі теоретичні показники, їхня реальна поведінка в комбінації з іншими криптографічними примітивами може суттєво відрізнятися. У науковій літературі ця проблема висвітлена

недостатньо, що актуалізує необхідність подальших досліджень, спрямованих на вивчення властивостей синтезованих S-блоків у практичних криптосистемах.

Метою цієї роботи є дослідження ефективності заміни стандартних S-блоків на синтезовані S-блоки з підвищеними криптографічними характеристиками, а також оцінка впливу таких замін на загальну стійкість та якість криптографічних алгоритмів.

Методи оцінки. Як було зазначено раніше, рівень конфузії та дифузії безпосередньо залежить від характеристик криптографічного примітиву, що входить до складу конкретного алгоритму. S-блок розглядається як один із ключових елементів сучасних криптографічних систем, оскільки саме завдяки йому забезпечується стійкість алгоритмів до криптоаналітичних атак, зокрема лінійного, диференційного та кореляційного аналізу. Таким чином, пошук та синтез високоякісних S-блоків залишається однією з центральних задач у дослідженнях криптографічних примітивів.

Щоб S-блок вважався криптографічно якісним, він має відповідати ряду наведених раніше критеріїв. Ці критерії оперують декомпозицією S-блоку на компонентні булеві функції $S = \{f_i\}, i = 1, 2, ..., k$, де k – відображає кількість входів/виходів S-блоку. Після розкладання S-блоку на компонентні булеві функції відбувається його оцінка відповідно до кожного цільового критерію.

Нелінійність. Відстань нелінійності – це ключове поняття, що визначається у часовій області як мінімальна відстань Геммінга між компонентними булевими функціями S-блока та всіма кодовими словами афінного коду і описується рівнянням

$$N_S = \min_i \left\{ \text{dist}(f_i, \varphi_j) \right\}, \; i = 1, ..., k, \; j = 1, ..., 2^{k+1}, \tag{1}$$

де f_i – i-та компонентна булева функція;

φ_j – кодове слово афінного коду;

dist(•) – оператор знаходження відстані Геммінга.

Також, існує альтернативний варіант оцінки за допомогою перетворення Уолша-Адамара, що визначається як

$$N_f = 2^{k-1} - \frac{1}{2} \max \left\{ \left| W(\omega) \right| \right\}, \tag{2}$$

де $\left| W(\omega) \right|$ – вектор трансформант перетворення Уолша-Адамара булевої функції f_i, що визначається як добуток таблиці істинності, представленої в експоненційній формі та матриці Уолша-Адамара

$$W = FA_N, \tag{3}$$

де матриця A_N представлена як

$$A_{2^k} = \begin{bmatrix} H_{2^{k-1}} & H_{2^{k-1}} \\ H_{2^{k-1}} & -H_{2^{k-1}} \end{bmatrix}, A_1 = 1. \tag{4}$$

Нелінійність всього S-блока визначається за найменшим визначеним значенням нелінійності серед усіх компонентних булевих функцій S-блока $N_S = \min_i \{N_{f_i}\}$. Цей принцип є дійсним як при обрахунках в часовій області (1), так і за допомогою перетворення Уолша-Адамара (2) [5].

Суворий лавинний критерій (SAC). Даний критерій оцінювання заснований на оцінці критерія поширення помилки [6]. Для його оцінки аналізуються мінімальні та максимальні значення ваг похідних компонентних булевих функцій

$$D_u f(x) = f(x) \oplus f(x \oplus u), \tag{5}$$

за усіма напрямками $\forall u \in V_k, wt(u) = 1$, де V_k – лінійний векторний простір векторів довжини *k,* а *wt*(□) – це оператор знаходження ваги Геммінга. Для того, щоб лавинні властивості були максимальними, усі похідні мають бути збалансованими, тобто їх вага

має дорівнювати $N/2$, що в свою чергу забезпечує вірогідність зміни виходу при зміні будь-якого входу рівною 0.5.

Критерій незалежності бітів (BIC). Відповідає за оцінку кореляції між компонентними булевими функціями S-блоку. Умовою виконання є наступне твердження.

Твердження 1. Якщо для двох будь-яких взятих компонентних булевих функцій f_a та f_b, де $a \neq b, 1 \leq a,b \leq k$ нова функція $f_a \oplus f_b$, що була утворена їхньою суперпозицією, має високу нелінійність та задовольняє суровому лавинному критерію, то S-блок, до складу якого входять ці функції f_a та f_b, вважається таким, що відповідає критерію незалежності бітів.

Тобто, для того, щоб цей критерій виконувався, компонентні булеві функції мають бути незалежними одна від одної, що гарантує, що при зміні одного вхідного біта, у вихідних бітах відбуваються максимально непередбачувані зміни [7].

Ймовірність лінійної апроксимації (LAP) – це критерій, що характеризує ймовірність лінійного наближення виходів S-блоку до заданих комбінацій його входів. Іншими словами, LAP визначає, наскільки легко виходи S-блоку можна передбачити за лінійною комбінацією входів. Чим слабкіший S-блок, тим вищим буде значення LAP, і тим вразливішою буде конструкцію до атак лінійного криптоаналізу. Для обчислення LAP використовується формула

$$LP_S = \frac{\max\limits_{\alpha,\beta\neq 0}(\#\{x \mid 0 \leq x < 2^k, \bigoplus\limits_{s=1}^{k} x[s] \bullet \alpha[s] = \bigoplus\limits_{t=1}^{k} S(x)[t] \bullet \beta[]\}) - 2^{k-1}}{2^k}, \quad (6)$$

де α, β – вхідна ти вихідна послідовності, відповідно, $[s]$ – позначає виділення s-го біта, а символ $\bullet$ позначає логічну операцію «І» («AND») [8].

Ймовірність диференціальної апроксимації (DAP) оцінює ймовірність того, що заданий вхідний диференціал S-блоку призведе до конкретного вихідного диференціалу протягом визначеної кількості раундів. Для її обчислення здійснюється повний перебір усіх можливих комбінацій вхідних і вихідних диференціалів для заданої кількості раундів. Підраховуються випадки появи кожного вихідного диференціалу, а DAP визначається як відношення кількості випадків бажаного вихідного диференціалу до загальної кількості перевірених пар вхідних/вихідних значень.

Чим нижче значення DAP, тим більш стійким є S-блок до диференціального криптоаналізу. DAP обчислюється за наступною формулою

$$DP(\mathrm{D}u, \mathrm{D}v) = \frac{\left|\{u \,\mathrm{O} V_k \mid S(u) \,\mathrm{E}\, S(u \,\mathrm{E}\, \mathrm{D}u) = \mathrm{D}v\}\right|}{2^k}, \quad (7)$$

де $\mathrm{D}u$ та $\mathrm{D}v$ – це вхідні і вихідні диференціали, відповідно.

Кореляційна незалежність вхідних і вихідних векторів S-блоку. Для оцінки даного критерію використовується максимальне значення серед модулів коефіцієнтів кореляції $\max\left\{\left|r_{i,j}\right|\right\}$ кореляційної матриці $R = \left\|r_{i,j}\right\|$, що визначає ступінь лінійного зв'язку між вихідним вектором y та вхідним вектором x. При цьому елементи $r_{i,j}$ кореляційної матриці R визначаються як

$$r_{i,j} = 1 - \frac{\sum\limits_{m=1}^{N}(x_{m,i} \oplus y_{m,j})}{N/2}, i, j = 0, \ldots, k-1, \quad (8)$$

Нижчі показники $\max\left\{\left|r_{i,j}\right|\right\}$ відображають вищий рівень криптографічної якості S-блоку [10].

Подання S-блока у вигляді функцій багатозначної логіки. Як зазначалося раніше, навіть за високих показників якості S-блоку у класичному булевому представленні можливі випадки, коли при розгляді того ж S-блоку у вигляді q-значної

логіки його криптографічні характеристики значно знижуються, іноді на кілька порядків. У таких умовах S-блок залишається вразливим до криптоаналізу, що базується на математичному апараті *q*-логіки.

Визначення 1. Функція *q*-значної логіки від *k* змінних – це відображення виду

$$\{0,1,2,...,q-1\}^k \to \{0,1,2,...,q-1\}. \quad (9)$$

Функції багатозначної логіки є більш загальними математичними об'єктами ніж булеві функції. Наприклад, якщо позначити $q=2$, то тоді *Визначення 1* стає визначенням булевої функції.

Таким чином, наприклад, S-блоки з 4 входами, довжини $N=16$ можуть бути представлені чотирма компонентними булевими функціями або двома компонентними 4-функціями (табл. 1).

Таблиця 1.

Представлення S-блоку у вигляді компонентних булевих та 4-функцій

Q	4	7	2	14	1	13	8	11	15	12	6	10	5	9	3	0
f_{20}	0	1	0	0	1	1	0	1	1	0	0	0	1	1	1	0
f_{21}	0	1	1	1	0	0	0	1	1	0	1	1	0	0	1	0
f_{22}	1	1	0	1	0	1	0	0	1	1	1	0	1	0	0	0
f_{23}	0	0	0	1	0	1	1	1	1	1	0	1	0	1	0	0
f_{40}	0	3	2	2	1	1	0	3	3	0	2	2	1	1	3	0
f_{41}	1	1	0	3	0	3	2	2	3	3	1	2	1	2	0	0

Якщо ж брати до уваги S-блоки більш практичної довжини 256, що використовується в багатьох сучасних криптоалгоритмах, то вони можуть бути представлені за допомогою восьми компонентних булевих функцій, чотирьох компонентних 4-функцій або ж двох компонентних 16-функцій [11].

Також у роботі [11] представлено математичний апарат для обчислення SAC для *q*-функцій.

Визначення 2. Вагою $\varpi(u)$ *q*-значного вектору називається число його ненульових компонент.

Визначення 3. Похідна функції f за напрямком вектору u – це функція

$$D_u f(x) = f(x \underset{q}{\oplus} u) - f(x) \,(\mathrm{mod}\, q), \quad (10)$$

де $\underset{q}{\oplus}$ означає додавання за модулем *q*.

Визначення 4. Функція *q*-значної логіки $f(x)$ задовольняє критерію поширення помилки $PC(u)$ щодо вектору $u \in V_k$, якщо її похідна у напрямку u є збалансованою функцією, тобто $0,1,...,q-1$ приймаються з однаковими ймовірностями $p(D_u f(x) = i(\mathrm{mod}\, q)) = \frac{1}{q}$ для всіх $i = 0,1,...,q-1$. Іншими словами, $K^0 = K^1 = ... = K^{q-1}$, де K^i – кількість наборів змінних значень, для яких похідна набуває значення *i*. Функція *q*-значної логіки $f(x)$ задовольняє критерію поширення $PC(m)$ степеню *m*, якщо вона задовольняє критерію поширення $PC(u)$ по відношенню до всіх векторів *u* ваги $1 \le \varpi(u) \le m$.

Визначення 5. Функція *q*-значної логіки $f(x)$ задовольняє SAC, якщо вона задовольняє критерію поширення $PC(1)$ степеню 1.

Метод синтезу високоякісних S-блоків на основі недвійкових афінних перетворень. На сьогодні конструкція Ніберг є однією з найпоширеніших у практичній криптографії. Вона використовується у складі широко відомого криптоалгоритму AES і відзначається високою збалансованістю щодо основних критеріїв криптографічної якості, таких як

нелінійність, стійкість до диференціального та лінійного криптоаналізу, а також здатність забезпечувати ефективну реалізацію конфузії та дифузії. Завдяки цим характеристикам S-блоки конструкції Ніберг забезпечують надійний рівень безпеки сучасних симетричних шифрів.

S-блок генерується за допомогою обчислення мультиплікативно зворотної величини в скінченному полі Галуа GF(2^8)

$$y_i = x_1^{-1} \text{modd}(2, G(x)),\ G(x) = x^8 + x^4 + x^3 + x + 1, \quad (11)$$

при цьому прийнято, що 0 співставляється сам з собою. У табл. 2 ми наводимо S-блок, побудований на основі (11).

Таблиця 2.

Мультиплікативна інверсія для усіх 8-бітних чисел у GF(2^8)

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**A**	**B**	**C**	**D**	**E**	**F**
0	0	1	141	246	203	82	123	209	232	79	41	192	176	225	229	199
1	116	180	170	75	153	43	96	95	88	63	253	204	255	64	238	178
2	58	110	90	241	85	77	168	201	193	10	152	21	48	68	162	194
3	44	69	146	108	243	57	102	66	242	53	32	111	119	187	89	25
4	29	254	55	103	45	49	245	105	167	100	171	19	84	37	233	9
5	237	92	5	202	76	36	135	191	24	62	34	240	81	236	97	23
6	22	94	175	211	73	166	54	67	244	71	145	223	51	147	33	59
7	121	183	151	133	16	181	186	60	182	112	208	6	161	250	129	130
8	131	126	127	128	150	115	190	86	155	158	149	217	247	2	185	164
9	222	106	50	109	216	138	132	114	42	20	159	136	249	220	137	154
A	251	124	46	195	143	184	101	72	38	200	18	74	206	231	210	98
B	12	224	31	239	17	117	120	113	165	142	118	61	189	188	134	87
C	11	40	47	163	218	212	228	15	169	39	83	4	27	252	172	230
D	122	7	174	99	197	219	226	234	148	139	196	213	157	248	144	107
E	177	13	214	235	198	14	207	173	8	78	215	227	93	80	30	179
F	91	35	56	52	104	70	3	140	221	156	125	160	205	26	65	28

Після обчислення мультиплікативної інверсії стає можливим виконати наступну трансформацію, передбачену конструкцією Ніберг – двійкове афінне перетворення. Цей етап є ключовим для формування необхідних криптографічних властивостей S-блоку, зокрема забезпечення стійкості до диференціального та лінійного аналізу

$$S_i = Ry_i + C \bmod 2, \quad (12)$$

де R – матриця афінного перетворення,

C – вектор-константа,

S_i – вихідні значення блоку AES.

Алгоритм AES пропонує наступну матрицю R і вектор-константу C

$$R = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}. \quad (13)$$

У роботі [12] запропоновано новий підхід до побудови криптографічних примітивів, що ґрунтується на використанні афінних перетворень у просторі багатозначної логіки. На відміну від класичного бінарного підходу, застосування q-функцій дозволяє отримати розширене представлення S-блоків та сформувати широкий клас їхніх модифікацій. Запропонований метод забезпечує синтез S-блоків вищої якості, які демонструють покращені криптографічні характеристики та підвищену стійкість до лінійних, диференціальних і кореляційних атак, що безпосередньо сприяє вдосконаленню сучасних симетричних шифрів.

У поєднанні з S-блоками [13], які задовольняють умовам суворого лавинного критерію (SAC) як для компонентних булевих функцій, так і для 4-функцій, ця схема дозволяє створювати S-блоки, що характеризуються високою криптографічною якістю. Згідно з запропонованим методом, афінне перетворення S-блоків виконується наступним чином

$$\alpha_j.S_i = \sum_{u=1}^{n} \alpha_j \times y_i R'(j,u) + C \bmod 4, \tag{14}$$

де позначення $\alpha_j.S_i$ означає взяття j-ї четвіркової цифри i-го елемента S-блока, тоді як позначення $R'(j,k)$ означає вилучення елемента з індексами (j,k) з матриці R', а $\times$ означає множення в полі Галуа $GF(4)$. В якості матриць R' використовуються модифіковані конструкції $C_{4,2}, C_{4,3}, C_{4,4}$

$$C_{4,21} = \begin{pmatrix} x & 0 & 0 & x_1 \\ 0 & 0 & x & x_1 \\ 0 & x & 0 & x_1 \\ 0 & 0 & 0 & 0 \end{pmatrix}; C_{4,22} = \begin{pmatrix} 0 & 0 & 0 & x_1 \\ x & 0 & x & x_1 \\ 0 & x & 0 & x_1 \\ 0 & 0 & 0 & 0 \end{pmatrix}; C_{4,23} = \begin{pmatrix} 0 & 0 & 0 & x_1 \\ 0 & 0 & x & x_1 \\ x & x & 0 & x_1 \\ 0 & 0 & 0 & 0 \end{pmatrix}; C_{4,24} = \begin{pmatrix} 0 & 0 & 0 & x_1 \\ 0 & 0 & x & x_1 \\ 0 & x & 0 & x_1 \\ x & 0 & 0 & 0 \end{pmatrix};$$

$$C_{4,31} = \begin{pmatrix} x & x_1 & 0 & 0 \\ 0 & x_1 & x & 0 \\ 0 & x_1 & 0 & x \\ 0 & 0 & 0 & 0 \end{pmatrix}; C_{4,32} = \begin{pmatrix} 0 & x_1 & 0 & 0 \\ x & x_1 & x & 0 \\ 0 & x_1 & 0 & x \\ 0 & 0 & 0 & 0 \end{pmatrix}; C_{4,33} = \begin{pmatrix} 0 & x_1 & 0 & 0 \\ 0 & x_1 & x & 0 \\ x & x_1 & 0 & x \\ 0 & 0 & 0 & 0 \end{pmatrix}; C_{4,34} = \begin{pmatrix} 0 & x_1 & 0 & 0 \\ 0 & x_1 & x & 0 \\ 0 & x_1 & 0 & x \\ x & 0 & 0 & 0 \end{pmatrix}; \tag{14}$$

$$C_{4,41} = \begin{pmatrix} x & 0 & x_1 & 0 \\ 0 & 0 & x_1 & x \\ 0 & x & x_1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}; C_{4,42} = \begin{pmatrix} 0 & 0 & x_1 & 0 \\ x & 0 & x_1 & x \\ 0 & x & x_1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}; C_{4,43} = \begin{pmatrix} 0 & 0 & x_1 & 0 \\ 0 & 0 & x_1 & x \\ x & x & x_1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}; C_{4,44} = \begin{pmatrix} 0 & 0 & x_1 & 0 \\ 0 & 0 & x_1 & x \\ 0 & x & x_1 & 0 \\ x & 0 & 0 & 0 \end{pmatrix}.$$

У виразі (14) x – це константне значення, що може одночасно набувати значення 1, 2, 3 для всіх x, а x_1 – це змінні, що які утворюють усі можливі комбінації чисел 1, 2 та 3.

Наслідком такого розмаїття є те, що з одного базового S-блоку можна синтезувати 7776 нових, так званих «дочірніх» конструкцій. Використовуючи різні перестановки, комбінації або модифікації вихідної матриці, можна генерувати великі множини унікальних, але споріднених S-блоків. Важливо, що серед отриманих конструкцій існують S-блоки з вищими криптографічними характеристиками, що забезпечує можливість створення сімейств шифрів або динамічної генерації примітивів з підвищеною стійкістю.

Це відкриває шлях до формування нових S-блоків, подібних до наведеного у табл. 3, кожен з яких може мати унікальні властивості, необхідні для конкретного криптографічного застосунку.

Таблиця 3.

S-блок конструкції Ніберг створений за допомогою афінного перетворення на базі 4-логіки

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**A**	**B**	**C**	**D**	**E**	**F**
0	15	6	47	128	178	210	121	131	245	166	107	93	29	216	104	244
1	46	59	164	43	198	81	167	150	190	115	73	221	192	112	220	9
2	19	188	51	58	69	191	142	238	87	97	201	154	84	200	33	228
3	152	23	18	7	155	170	250	109	117	225	130	79	224	53	252	76
4	135	173	194	141	83	56	193	243	231	41	28	116	90	42	118	156
5	185	230	169	175	16	229	215	71	49	8	92	255	50	94	136	66
6	254	177	187	145	253	207	111	4	32	72	219	21	74	160	102	22
7	149	147	133	218	235	123	44	217	96	195	13	52	180	126	14	98
8	236	67	204	197	186	64	113	17	168	158	54	101	171	55	222	27
9	103	232	237	248	100	85	5	146	138	30	125	176	31	202	3	179
A	240	249	208	127	77	45	134	124	10	89	148	162	226	39	151	11
B	209	196	91	212	57	174	88	105	65	140	182	34	63	143	35	246
C	1	78	68	110	2	48	144	251	223	183	36	234	181	95	153	233
D	106	108	122	37	20	132	211	38	159	60	242	203	75	129	241	157
E	120	82	61	114	172	199	62	12	24	214	227	139	165	213	137	99
F	70	25	86	80	239	26	40	184	206	247	163	0	205	161	119	189

Імплементація та оцінка стійкості сучасних шифрів із високоякісними S-блоками. Для оцінки криптографічних якостей S-блоків, синтезованих у [12] було використано наступні блокові симетричні шифри: AES, Camelia, SM-4, Kalyna та ARIA.

AES. Це симетричний блоковий шифр, що набув широкого застосування, що працює з блоками фіксованого розміру і підтримує три довжини ключа: 128, 192 та 256. Завдяки своїй високій швидкості та стійкості, AES став міжнародно визнаним алгоритмом і широко використовується у різних сферах, включаючи Wi-Fi, VPN та навіть апаратно підтримується у багатьох сучасних пристроях [14].

Camelia. Шифр, що був розроблений спільно японськими компаніями Mitsubishi та NTT. Як і AES підтримує три довжини блоку: 128, 192 та 256. Ключова особливість – це використання додаткових FL-функцій, що застосовуються кожні 6 раундів для підвищення криптографічної стійкості. Фактично, Camelia була визнана міжнародними стандартами ISO/IEC і стала досить популярною в Японії, фактично є одним з японських стандартів [15].

Калина. Український стандарт, що набув чинності у 2014 році. Він підтримує довжину блоку і ключа від 128 біт до 512 біт, а також має високий рівень криптографічної стійкості з достатнім запасом у разі винайдення або створення нових атак протягом тривалого часу. Також, на відміну від AES, в нього значно збільшена кількість циклів шифрування, а також застосована принципово нова схема створення підключів [16].

SM-4. Китайський стандарт шифрування, що оперує блоками даних розміром 128 біт і використовує ключ відповідної довжини. Для шифрування він використовує 32 раунди з операціями XOR, використанням S-блоків та циклічних зсувів [17].

ARIA. Південно-корейський стандарт, що був затверджений у 2004 році. Для шифрування використовує блоки розмірів від 128 біт до 256 біт. Як і AES, ARIA створена на базі SP-мережі (підстановочно-перестановочної мережі), що включає у себе декілька раундів [18].

Для початку експерименту було проведено початкову оцінку S-блоків, що входять до складу наведених криптоалгоритмів та цільових S-блоків з [12], що будуть використовуватись у якості альтернативи (табл. 4).

Таблиця 4.

Порівняння криптографічних властивостей S-блоків провідних криптографічних стандартів

S-box	Nonlinearity	SAC	BIC Nonlinearity	BIC SAC	LAP	DAP	$\max\left\{\left\|r_{i,j}\right\|\right\}$	SAC of component 4-functions
AES	112	0.5032	112	0.5057	0.0625	0.0156	0.125	0.6484
Camelia	112	0.5469	112	0.5033	0.0625	0.0156	0.1094	0.6563
	112	0.5469	112	0.5033	0.0625	0.0156	0.1094	0.6172
	112	0.5469	112	0.5033	0.0625	0.0156	0.1094	0.6172
	112	0.5469	112	0.5033	0.0625	0.0156	0.1094	0.6172
Калина	104	0.5625	106.6429	0.5015	0.0938	0.0313	0.1563	0.6094
	104	0.5938	106.8571	0.5024	0.0938	0.0313	0.1719	0.6250
	104	0.6094	107.0714	0.6250	0.0938	0.0313	0.1563	0.6250
	104	0.5781	107.0714	0.5024	0.0938	0.0313	0.1563	0.6254
ARIA	112	0.5625	112	0.5046	0.0625	0.0156	0.125	0.6016
	112	0.5625	112	0.5030	0.0625	0.0156	0.0938	0.5781
SM-4	112	0.5625	112	0.5049	0.0625	0.0156	0.125	0.6484
Affine S-boxes[13]	96	0.5	77.7143	0.4799	0.5	1	0	0.5
	96	0.5	77.7143	0.5045	0.5	1	0	0.5
	96	0.5	77.7143	0.5022	0.5	1	0	0.5
	96	0.5	77.7143	0.4799	0.5	1	0	0.5

Як видно з табл. 4, вихідні S-блоки провідних криптоалгоритмів демонструють високі криптографічні характеристики, що й обумовлює їхню популярність і широке застосування. Проте для оцінки ефективності нових конструкцій було перевірено,

наскільки якіснішими стають криптоалгоритми після заміни оригінальних S-блоків на синтезовані у [12]. Для оцінки ефективності застосовувалися стандартні стохастичні тести NIST (NIST Statistical Test Suite), які широко використовуються для перевірки криптографічних алгоритмів та генераторів випадкових послідовностей. Ці тести дозволяють визначити, наскільки вихідні дані після шифрування відповідають властивостям випадковості, що безпосередньо пов'язано з рівнем криптографічної стійкості. Іншими словами, якщо зашифровані дані демонструють статистичні властивості випадкових послідовностей, алгоритм вважається більш стійким до криптоаналітичних атак.

Таблиця 5.

Порівняння ефективності оригінальних та модифікованих криптоалгоритмів

Криптоалгоритм	Пройдено тестів (Оригінал)	Пройдено тестів (Модифікація)
AES	94	96
ARIA	95	96
Kalyna	99	99
SM4	94	97
Camelia	92	99

Відповідно до результатів, наведених у табл. 5, S-блоки, синтезовані у [12], демонструють високий рівень криптографічних характеристик у різних обчислювальних системах. Хоча в булевому сенсі їх показники можуть бути дещо нижчими порівняно з окремими альтернативними S-блоками, їх впровадження суттєво підвищує загальну криптографічну стійкість алгоритмів. Це свідчить про те, що дані S-блоки є ефективним інструментом для оптимізації існуючих криптографічних конструкцій, а також можуть бути рекомендовані для розробки нових стійких криптоалгоритмів.

Висновки. Результати дослідження показали, що інтеграція синтезованих високоякісних S-блоків забезпечує потрійний ефект: високу якість у традиційних метриках булевих функцій, значні характеристики у метриках функцій багатозначної логіки та практичне підвищення криптографічної стійкості алгоритмів при їх вбудовуванні, що раніше залишалося проблемою навіть для провідних світових стандартів.

Отримані результати підтверджують, що навіть за умов, коли синтезовані примітиви мають дещо нижчі показники у класичному булевому сенсі, комплексна збалансованість їх криптографічних властивостей у кількох системах представлення забезпечує підвищену криптографічну якість шифру в цілому. Це безпосередньо відобразилося на результатах стохастичних тестів NIST: у більшості випадків модифіковані алгоритми перевищили оригінальні за кількістю успішно пройдених тестів, демонструючи більшу випадковість та непередбачуваність шифрованих даних.

Таким чином, дослідження доводить перспективність розробленого підходу для побудови нового покоління симетричних шифрів. Синтезовані S-блоки можуть стати основою динамічних криптографічних примітивів, що дозволяють створювати адаптивні та самозахисні криптосистеми, стійкі до широкого спектра атак, включно з тими, що ґрунтуються на функціях багатозначної логіки.

У практичному вимірі це означає можливість створення більш захищених протоколів комунікації, стійких до квантових атак, а також криптографічних рішень для IoT, хмарних сервісів та критичної інфраструктури, де вимоги до інформаційної безпеки зростають у геометричній прогресії.

Отже, результати цієї роботи не лише підсумовують успішність запропонованого методу, а й відкривають шлях до формування нового напряму розвитку сучасної симетричної криптографії, заснованого на синтезі примітивів із стійкістю у сенсі функцій багатозначної логіки.

Список літератури

1. Hussain U. A Comparative Survey of Symmetric and Asymmetric Key Cryptography Algorithms. *2nd International Multidisciplinary Conference on Emerging Trends in Engineering Technology*. *Khairpur*. 2024. P. 257-262.
2. Thakor V. A., Razzaque M. A., Khandaker M. R.A. Lightweight Cryptography Algorithms for Resource-Constrained IoT Devices: A Review, Comparison and Research Opportunities. *IEEE Access*. 2021. Vol. 9. P. 28177–28193. doi: 10.1109/access.2021.3052867
3. Sokolov A., Kazakova N., Kuzmenko L., Mahomedova M. Prerequisites for developing a methodology for estimating and increasing cryptographic strength based on many-valued logic functions. *CEUR Workshop Proceedings*. 2021. 2923. P. 107-116. URL: https://ceur-ws.org/Vol-2923/paper12.pdf
4. Baigneres T., Stern J., Vaudenay S. Linear cryptanalysis of non binary ciphers. International Workshop on Selected Areas in Cryptography. Springer, Berlin, Heidelberg, 2007. P. 184-211.
5. Zahid A. H. et al. Efficient Dynamic S-Box Generation Using Linear Trigonometric Transformation for Security Applications. *IEEE Access*. 2021. Vol. 9. P. 98460-98475. DOI: 10.1109/access.2021.3095618
6. Banga A. ChessCrypt: enhancing wireless communication security in smart cities through dynamically generated S-Box with chess-based nonlinearity. *Scientific Reports*. 2024. Vol. 14, No. 1. P. 1-25. DOI: 10.1038/s41598-024-77927-0
7. Jamal S. S. Secure S-box construction with 1D chaotic maps and finite field theory for block cipher encryption. *Alexandria Engineering Journal*. 2025. Vol. 125. P. 278–296. DOI: 10.1016/j.aej.2025.03.109
8. Kazakova N. Nonlinearity of Many-Valued Logic Component Functions of Modern Cryptographic Algorithms S-boxes. *Procedia Computer Science*. 2021. Vol. 192. P. 2731-2741. DOI: 10.1016/j.procs.2021.09.043
9. Farah M. A. B., Farah A., Farah T. An image encryption scheme based on a new hybrid chaotic map and optimized substitution box. *Nonlinear Dynamics*. 2019. Vol. 99, No. 1. P. 1-24. DOI: 10.1007/s11071-019-05413-8
10. Sokolov A.V., Zhdanov O.N. Prospects for the Application of Many-Valued Logic Functions in Cryptography. *International Conference on Theory and Applications of Fuzzy Systems and Soft Computing, Springer, Cham*. 2018. P. 331-339. DOI: 10.1007/978-3-319-91008-6_33
11. Sokolov A.V., Radush V.V. A method for synthesis of S-boxes with good avalanche characteristics of component Boolean and quaternary functions. *Journal of Discrete Mathematical Sciences and Cryptography*. 2022. Vol. 26. Is. 2. P. 1-12. DOI: 10.1080/09720529.2021.1964727
12. Karpinski M.. Development of High-Quality Cryptographic Constructions Based on Many-Valued Logic Affine Transformations. *Electronics.* 2025. Vol. 14, No. 10. P. 1-22. DOI: 10.3390/electronics14102094
13. Sokolov A.V., Radush V.V. The method for synthesis of high-quality S-boxes based on many-valued logic functions. *Informatics and mathematical methods in simulation*. 2022. Vol. 12, No. 3. P. 219-225. DOI: 10.15276/imms.v12.no3.219
14. Jang K. Quantum Analysis of AES / IACR Communications in Cryptology. 2025. Vol. 2, No. 1. P. 1-57. DOI: 10.62056/ay11zo-3y
15. The Camellia Cipher. Security and So Many Things. URL: https://asecuritysite.com/blog/2023-11-18_The-Camellia-Cipher-44d2d044de4d.html.
16. Єфіменко А. А., Байлюк Є. М., Покотило О. А. Порівняльний аналіз алгоритму симетричного блокового перетворення «Калина» (ДСТУ 7624:2014) з іншими міжнародними стандартами шифрування даних. *Збірник наукових праць ЖВІ*. 2018. № 15. С. 156-162.

17. Бондаренко О., Філобок Є., Козіна Г. Реалізація алгоритму блочного шифрування SM4. інформаційні технології: теорія і практика. *ІІІ Всеукр. науково-практ. Інтернет-конф. здобувачів вищ. освіти і молодих уч.* 2025 р. С. 36-37.
18. RFC 5794: A Description of the ARIA Encryption Algorithm. RFC Editor. URL: https://www.rfc-editor.org/rfc/rfc5794.html.

IMPROVING THE RESISTANCE OF MODERN BLOCK CIPHERS USING HIGH-QUALITY S-BOXES

V.V. Radush

National Odesa Polytechnic University
1, Shevchenko Ave., Odesa, 65044, Ukraine
Email: radush9860@gmail.com

Ensuring a high level of cryptographic stability of modern symmetric ciphers directly depends on the quality of the S-boxes used in them. Traditionally, the analysis of S-boxes is limited to the apparatus of Boolean functions; however, recent research shows that under the conditions of attacks based on many-valued logic, some standard S-boxes may demonstrate an insufficient level of cryptographic stability. This creates the need for a comprehensive assessment of such cryptographic primitives to ensure their quality in both Boolean and many-valued representations. An experimental evaluation of the quality of S-boxes used in modern symmetric block ciphers was performed, with an emphasis on practical efficiency in real encryption and assessment of stochastic properties of the cryptogram. In order to strengthen the cryptographic characteristics of the ciphers, instead of the original substitutions, the author's S-boxes were used, built based on non-binary affine transformations, which demonstrate high cryptographic properties when represented by both Boolean functions and many-valued logic functions. The methodology included two complementary stages: calculation of key S-box quality metrics – nonlinearity, fulfillment of the strict avalanche criterion (SAC), bit independence criterion (BIC), linear (LAP) and differential (DAP) approximation probability; practical verification by encrypting information arrays with modified ciphers and subsequent statistical analysis of the original cryptograms using NIST tests. The research covered AES, Camellia, Kalyna, SM4, and ARIA cryptographic algorithms. The calculated indicators, together with stochastic verification, showed that S-boxes based on quaternary affine transformations provide increased balance and stability. In most cases, the modified versions of the ciphers outperformed the originals in the number of NIST tests passed (in particular, AES, ARIA, SM4, Camellia), which indicates an increase in the randomness and cryptographic resistance of the original sequences. It is important to emphasize the experimental nature of the research: a comparative and practical verification of S-boxes based on affine transformations was performed. The obtained data indicate the prospects for further integration of such primitives into communication protocols, protection of IoT devices, cloud services, and critical infrastructure, where the requirements for randomness and stability are particularly high. The combination of theoretical quality criteria and practical verification of cryptograms allows us to speak of a reliable methodological basis for making decisions on the implementation of such S-boxes in real systems.

Keywords: cryptographic primitives; block ciphers; S-boxes; affine transformations; many-valued logic; nonlinearity; avalanche criterion; NIST Statistical Test Suite; cryptanalysis; cryptogram randomness.

DOI 10.15276/imms.v15.no4.625
УДК 004.056.5

Informatics and Mathematical Methods in Simulation
Vol.15 (2025), No. 4, pp. 625-635

ДОСЛІДЖЕННЯ СТІЙКОСТІ ТРАНСФОРМАНТ ПЕРЕТВОРЕННЯ УОЛША-АДАМАРА ДО СТИСНЕННЯ MPEG3 В ЗАДАЧАХ АУДІОСТЕГАНОГРАФІЇ

[1]А. В. Соколов, [2]М. В. Хименко

[1]Національний університет «Одеська юридична академія»
23, Фонтанська дорога, Одеса, 65009, Україна
[2]Національний університет «Одеська політехніка»
1, Шевченка пр., Одеса, 65044, Україна

Представлено новий напрямок розвитку аудіостеганографії – адаптація та обґрунтування методів кодового управління для підвищення стійкості прихованих повідомлень в аудіоконтейнерах до сучасних атак стисненням, зокрема до MPEG3. Автори проводять системний аналіз сучасних підходів до аудіостеганографії та підкреслюють їхні обмеження – втрату прихованих даних під час стиснення або надмірну вимогливість до обчислювальних ресурсів, що ускладнює використання в мобільних та вбудованих системах. У цьому контексті пропонується адаптувати до аудіоконтейнерів концепцію кодового управління, яка вже продемонструвала високу ефективність для зображень. Такий підхід відкриває можливість здійснювати контрольоване вбудовування даних у стійкі компоненти сигналу, поєднуючи непомітність, стійкість до атак і обчислювальну ефективність. Ключовим внеском роботи є розробка алгоритму ідентифікації трансформант Уолша-Адамара, які мінімально спотворюються під час стиснення MPEG3. Алгоритм послідовно порівнює пари блоків вихідного (FLAC) і стисненого (MPEG3) сигналів, обчислює перетворення Уолша-Адамара, накопичує статистики змін та візуалізує частоти мінімального спотворення трансформант Уолша-Адамара у вигляді гістограм, що дозволяє об'єктивно виділяти пріоритетні позиції для вбудовування. Експерименти виконані на широкому наборі HiFi-аудіозаписів із тестуванням кількох бітрейтів (320, 256, 192, 128 кбіт/с) та різних довжин блоків, що забезпечує репрезентативність та надійність висновків. Результати експериментів демонструють стійку закономірність: трансформанти з індексами 8 і 12 стабільно виявляють найвищу стійкість до спотворень при стисканні і є пріоритетними для вбудовування прихованих даних; при зниженні бітрейту до 192 і 128 кбіт/с окрім зазначених трансформант практично допустимо (і доцільно в ряді сценаріїв) використання трансформант 0 і 4, що розширює набір безпечних позицій для кодового управління. Крім того, при збільшенні довжини блоку до 64 елементів, показано, що стійкою є трансформанта 32. Практична значимість дослідження полягає в тому, що виділені стійкі трансформанти можуть стати основою для створення адаптивних, високоефективних алгоритмів стеганографії з кодовим управлінням: вони дозволяють поєднувати невидимість, стійкість до поширених алгоритмів стиску (включаючи MPEG3) і низькі обчислювальні витрати, що робить запропонований підхід придатним для впровадження в реальні системи. Робота також відкриває перспективи подальших досліджень – розширення аналізу на інші алгоритми стиснення та мультимедійні формати, інтеграцію методів машинного навчання для автоматичної адаптації кодових слів та розробку нових стеганографічних схем.

Ключові слова: аудіостеганографія, кодове управління, перетворення Уолша-Адамара, стеганоповідомлення, стійкість до стиснення, MPEG3-компресія, трансформанти, інформаційна безпека.

Вступ і постановка задачі. Сучасні інформаційні системи дедалі частіше стають об’єктом атак, спрямованих на несанкціоноване отримання, модифікацію чи блокування даних. У таких умовах особливого значення набувають системи стеганографічного захисту інформації. На відміну від криптографії, що забезпечує конфіденційність вмісту повідомлення, стеганографія дозволяє приховати сам факт його існування, що робить її

потужним інструментом у сфері кібербезпеки, цифрових комунікацій та захисту авторських прав [1].

Актуальність стеганографії обумовлена зростанням обсягів мультимедійного контенту, який використовується як контейнер для прихованих повідомлень. Аудіо- та відеофайли, завдяки своїй надмірності та складній структурі, відкривають широкі можливості для вбудовування даних без помітного погіршення надійності сприйняття. Разом із тим, розвиток алгоритмів стиснення з втратами висуває нові виклики, оскільки такі перетворення можуть спотворювати або знищувати стеганографічні вбудовування. Тому дослідження стійкості методів стеганографії до сучасних форматів стиснення є одним із ключових напрямів наукових пошуків сьогодення.

У науковій літературі значна частка досліджень присвячена стеганографії у цифрових зображеннях, що пояснюється їхньою поширеністю та зручністю для вбудовування даних [2]. Проте аудіоконтейнери також відіграють вагому роль у сфері прихованого передавання інформації. Аудіо широко використовується в телекомунікаціях, медіаіндустрії та потокових сервісах, тому вдосконалення методів стеганографії для звукових сигналів є вкрай актуальним. Розробка стійких до стиснення та інших видів атак алгоритмів приховування в аудіо відкриває перспективи для створення більш надійних і практично орієнтованих систем захисту інформації.

До сучасних стеганографічних методів висувається низка ключових вимог [3]. Насамперед це надійність сприйняття, тобто відсутність помітних змін у контейнері для людини чи стандартних засобів аналізу якості. Важливим критерієм є також стійкість до атак, спрямованих на пошкодження або знищення прихованого повідомлення, а також захищеність від стеганоаналізу, що забезпечує невиявність факту прихованого передавання даних. Сьогодні до цих класичних вимог додається ще один критично важливий аспект – обчислювальна ефективність. Умови використання стеганографії дедалі частіше передбачають реалізацію алгоритмів на пристроях із обмеженими ресурсами, зокрема в системах Інтернету речей та на мобільних платформах. Це вимагає створення методів, які поєднують високу стійкість і невиявність із мінімальними витратами обчислювальних ресурсів.

Наразі запропоновано чимало методів стеганографії для аудіо, від простих бітових модифікацій до сучасних підходів із застосуванням машинного навчання й методів «coverless».

У роботі [4] запропонований метод аудіостеганографії, що базується на модифікації окремих бітів аудіосигналу. Хоча цей підхід простий у реалізації і може бути ефективним при низьких бітрейтах, він страждає від низької стійкості до різних методів аналізу та обробки сигналів. Найменші зміни в аудіофайлі можуть призвести до помітних спотворень, що робить його вразливим до атак і знижує його застосування в реальних умовах.

Запропонований метод coverless аудіостеганографії [5] використовує генеративні змагальні мережі (GAN) для синтезу стеганографічного аудіо без необхідності у вихідному аудіофайлі. Хоча це рішення підвищує скритність, воно вимагає значних обчислювальних ресурсів і може страждати від обмеженої здатності відновлення прихованої інформації. Крім того, якість синтезованого аудіо може змінюватись, що впливає на сприйняття кінцевого користувача.

Метод [6] пропонує сегментацію аудіофайлу на передній та задній плани для більш ефективного вбудовування прихованої інформації. Однак підхід може бути чутливим до різних типів аудіофайлів та умов запису. Крім того, алгоритм може вимагати попереднього аналізу та налаштування для кожного конкретного випадку, що обмежує його універсальність та зручність використання.

Метод [7] поєднує в собі придушення мікромодуляції амплітуди та використання узагальненої аудіо-внутрішньої енергії для підвищення стійкості до спотворень. Однак, незважаючи на поліпшення у скритності та стійкості, метод може бути складним у

реалізації та вимагати значних обчислювальних ресурсів. Крім того, ефективність методу може залежати від характеристик вихідного аудіофайлу, що обмежує його універсальність.

Підхід [8] використовує кластеризацію з диференціальною приватністю для створення стеганографічного аудіо без необхідності у вихідному файлі. Хоча метод забезпечує високий рівень конфіденційності та скритності, він може бути чутливим до якості та різноманітності вихідних даних. Крім того, складність алгоритму та вимоги до обчислювальних ресурсів можуть обмежувати його застосування у реальних умовах.

Як видно з проведеного аналізу представників сучасних стеганографічних методів, жоден із них не позбавлений суттєвих недоліків: одні вразливі до стандартних перетворень і стиснення (що призводить до втрати вбудованої інформації), інші потребують значних обчислювальних ресурсів або мають обмежену універсальність щодо жанру й формату аудіо, ще інші – вразливі до сучасних методів стеганоаналізу. Через це питання створення одночасно стійких, невиявних та ефективних у обчислювальному сенсі аудіостеганографічних методів залишається відкритим і потребує подальших досліджень.

У галузі стеганографії для цифрових зображень справжнім проривом став підхід на основі кодового управління [9]. Використання цієї концепції дозволило вивести стеганографічні методи на новий рівень, забезпечивши одночасне дотримання ключових вимог: надійності сприйняття, стійкості до атак та захищеності від стеганоаналізу. Крім того, алгоритми з кодовим управлінням відзначаються високою обчислювальною ефективністю, що робить їх придатними для практичного застосування навіть на пристроях з обмеженими обчислювальними ресурсами. Це відкриває перспективи поширення даного підходу і на інші мультимедійні контейнери, зокрема аудіосигнали.

Незважаючи на успіхи кодового управління у стеганографії для зображень, на сьогодні цей метод не адаптований для аудіоконтейнерів. Більше того, залишаються невивченими навіть базові характеристики застосування підходу до звукових сигналів. Зокрема, невідомо, які трансформанти аудіо могли б забезпечити найкращу роботу алгоритму, зберігаючи стійкість до атак, захищеність від стеганоаналізу та мінімальні обчислювальні витрати. Це обґрунтовує необхідність проведення фундаментальних досліджень для визначення оптимальних параметрів і стратегій застосування кодового управління у аудіостеганографії.

У рамках даного дослідження планується дослідити трансформанти аудіосигналів, отримані за допомогою перетворення Уолша-Адамара, з точки зору їхньої стійкості до атак проти прихованого повідомлення. Особлива увага буде приділена впливу стиснення аудіо, оскільки воно є однією з найпоширеніших форм модифікації сигналу, здатною суттєво спотворювати вбудовану інформацію. Результати цього аналізу дозволять визначити, які трансформанти є найбільш підходящими для подальшого застосування методів кодового управління в аудіостеганографії.

Метою цієї роботи є підвищення стійкості стегноповідомлень у аудіо контейнерах на основі кодового управління до атак стисненням шляхом дослідження трансформант перетворення Уолша-Адамара.

Для досягнення поставленої мети необхідно вирішити наступні задачі:

1. Дослідження модальності алгоритмів стиснення аудіо, які можуть виступати як атака на приховане повідомлення.
2. Дослідження вплив алгоритмів стиснення на трансформанти Уолша-Адамара аудіоконтейнерів.
3. Визначення трансформанти Уолша-Адамара, що демонструють найвищу стійкість до атак стисненням.

Розв’язання зазначених завдань закладе фундамент для подальшого застосування методів кодового управління в аудіоконтейнерах. Визначення стійких до стиснення трансформант Уолша-Адамара дозволить розробляти алгоритми приховування

інформації, що поєднують високу стійкість до атак, захищеність від стеганоаналізу та обчислювальну ефективність, необхідну для реалізації на мобільних та вбудованих пристроях. Таким чином, результати цього дослідження стануть основою для створення практично орієнтованих систем аудіостеганографії нового покоління.

Огляд концепції кодового управління та її застосування для аудіоконтейнерів. На сьогодні створена ціла плеяда методів стеганографії з кодовим управлінням. Серед них можна виділити класичний метод, заснований на бінарних кодових словах [9], метод на основі багаторівневих кодових слів [10,11], метод зі сліпим декодуванням [12], а також алгоритми, що забезпечують множинний доступ до прихованої інформації [13]. Кожен із цих підходів по-своєму забезпечує дотримання ключових вимог стеганографії, поєднуючи надійність сприйняття, стійкість до атак і захищеність від стеганоаналізу, при цьому демонструючи високу обчислювальну ефективність.

Основою всіх цих методів є перетворення Уолша-Адамара [14], яке визначається за допомогою наступного співвідношення

$$V = YH_N, \tag{1}$$

де матриця Адамара H_N порядку N задається за допомогою конструкції Сильвестра

$$H_{2^k} = \begin{bmatrix} H_{2^{k-1}} & H_{2^{k-1}} \\ H_{2^{k-1}} & -H_{2^{k-1}} \end{bmatrix}, \; H_1 = 1. \tag{2}$$

Основна ідея стеганографії з кодовим управлінням полягає у лінійності перетворення Уолша-Адамара, що дозволяє ефективно працювати з трансформантами сигналу, а також у адитивний спосіб вбудовувати приховане повідомлення. Розпишемо цю ідею застосовно до аудіоконтейнерів, для яких доречним є саме одновимірний варіант перетворення Уолша-Адамара. Нехай у черговий вектор блоку аудіоконтейнера Y_i необхідно вбудувати черговий біт інформації d_i, який представляється у вигляді знакового кодування кодового вектора C_i тоді вектор стеганоповідомлення матиме вигляд

$$S_i = Y_i + (-1)^{d_i} C_i. \tag{3}$$

Знаходячи перетворення Уолша-Адамара вектора S_i згідно до (1) та застосовуючи властивість лінійності цього перетворення отримуємо наступне співвідношення

$$V_{S_i} = (Y_i + (-1)^{d_i} C_i) H_N = Y_i H_N + (-1)^{d_i} C_i H_N. \tag{4}$$

Поданий вираз демонструє, що конкретний характер впливу на трансформанти перетворення Уолша-Адамара вектора блоку контейнера Y_i визначається видом кодового слова C_i, яке використовується під час вбудовування. Іншими словами, вибір кодового слова безпосередньо задає, які саме трансформанти будуть модифіковані та яким чином. Якщо ж кодове слово підібране таким чином, щоб воно цілеспрямовано впливало на задані трансформанти перетворення Уолша-Адамара, то з'являється можливість вбудовувати додаткову інформацію саме в ті компоненти, які є найбільш придатними з точки зору стійкості до атак чи невиявності. Це відкриває шлях до побудови адаптивних стеганографічних методів, де управління вбудовуванням здійснюється на рівні структури трансформант.

Методи, засновані на цій ідеї, вже продемонстрували свою ефективність у стеганографії для зображень. Зокрема, вони забезпечують високу стійкість до атак проти вбудованого повідомлення та до атак стеганоаналізу, при цьому вплив на елементи зображення залишається мінімальним і практично непомітним. Важливою перевагою є також те, що завдяки роботі у просторовій області такі методи характеризуються надзвичайно високою швидкодією, що робить їх придатними для використання в реальних системах із жорсткими обмеженнями на обчислювальні ресурси.

Алгоритми стиску аудіосигналів, які можуть застосовуватися для атак. Сучасні алгоритми стиснення аудіо базуються на принципах психоакустики, видаляючи з

сигналу компоненти, малопомітні або нечутні для людського слуху. Це дозволяє істотно зменшити обсяг даних при збереженні задовільної якості звучання. До найбільш відомих підходів належать AAC (Advanced Audio Coding) [15], що забезпечує підвищену ефективність кодування, та OGG Vorbis [16], який виступає відкритою альтернативою комерційним стандартам. Водночас у практичному використанні ключову роль продовжує відігравати алгоритм MPEG3 (MPEG-1 Audio Layer III) [17], який завдяки оптимальному співвідношенню між якістю відтворення та ступенем стиснення залишається найпоширенішим форматом аудіо сьогодні. На рис. 1 подано узагальнену схему роботи алгоритму MPEG3, яка відображає основні етапи обробки сигналу – від аналізу спектра та застосування психоакустичної моделі до квантування і кодування бітового потоку.

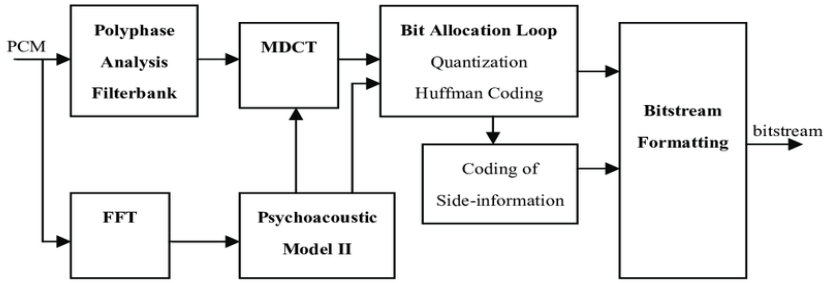


Рис. 1. Схема роботи MPEG3 кодера

На рис. 1 представлено стандартизовану схему роботи MPEG3-кодера, з якої можна виділити такі основні етапи перетворення lossless-файлу (файлу без втрат якості) у формат MPEG3:

– Розбиття, аналіз та квантування. На цьому етапі задається вихідна якість сигналу (як правило частота дискретизації становить 44.1 кГц) та точність квантування (16 біт), що визначають подальшу обробку звуку.

– Перетворення за допомогою MDCT. Вхідна послідовність ділиться на блоки (зазвичай по 576 семплів), після чого кожен блок піддається модифікованому косинусному перетворенню (MDCT).

– Швидке перетворення Фур'є (FFT – Fast Fourier Transform) [8]. Використовується для аналізу спектра та переходу між часовою та частотною областями представлення сигналу.

– Частотне маскування та психоакустичне моделювання [9]. На цьому кроці відкидаються частоти, нечутні для людини, а також ті, що неістотно впливають на якість сприйняття, що суттєво зменшує розмір файлу.

– Квантування та кодування за допомогою коду Гаффмана [10]. Частотні компоненти квантизуються залежно від їхньої важливості (на основі психоакустичних моделей), після чого застосовується кодування Гаффмана для подальшого зменшення обсягу даних.

Для формування вихідного MP3-файлу ключову роль відіграє модифіковане дискретне косинусне перетворення (MDCT). Воно застосовується до послідовності, що попередньо пройшла обробку за допомогою 32-смугового багатофазового квадратурного фільтра (PQF). Завдяки цьому поєднанню забезпечується перехід від часової області до частотної з урахуванням перекриття блоків, що дозволяє зменшити спотворення на межах сегментів та досягти високої якості відтворення при суттєвому зменшенні обсягу даних.

Вектор трансформант p_k перетворення MDCT [18] для заданого блоку $\{x_k\}$ задається як

$$p_k = \sum_{n=0}^{2N-1} x_n \cos\left(\frac{\pi}{N}\left(n+\frac{1}{2}+\frac{N}{2}\right)\left(k+\frac{1}{2}\right)\right). \quad (5)$$

Запропонований алгоритм дослідження. Для проведення експериментів було використано набір із 155 звукових фрагментів та музичних композицій у HiFi-якості у форматі FLAC [11]. Для оцінки ефективності стиснення застосовувався алгоритм з різними бітрейтами.

У більшості випадків стандартним вважається бітрейт 192 Кбіт/с, який забезпечує оптимальний баланс між якістю звучання та розміром вихідного файлу. При такому рівні стиснення звук зберігає повну розбірливість і не втрачає помітних деталей навіть при відтворенні на стандартних пристроях прослуховування.

За основу для пошуку було взято набір з 108 звуків та пісень HiFi якості у форматі FLAC. Алгоритм стискає вхідні файли з бітрейтами 320 Кбіт/с, 256 Кбіт/с, 192 Кбіт/с, 128 Кбіт/с.

Для демонстрації ефективності роботи алгоритму MP3-кодування було проведено експериментальне стиснення набору аудіофайлів у форматі FLAC із подальшим порівнянням їхніх розмірів у вихідному (нестисненому) вигляді та після перетворення з різними бітрейтами.

У табл. 1 наведено приклад значення розмірів набору файлів у мегабайтах для різних бітрейтів, що дозволяє оцінити співвідношення між ступенем стиснення обсягом пам’яті, який займатимуть аудіофайли.

Таблиця 1.

Залежність розміру вихідного набору даних від коефіцієнту збереження якості при стисненні

Бітрейт	Без стиснення (.flac)	320 Кбіт/с (.mp3)	256 Кбіт/с (.mp3)	192 Кбіт/с (.mp3)	128 Кбіт/с (.mp3)
Розмір	2.92 ГБ	1.2 ГБ	1.03 ГБ	878 МБ	699 МБ

Як можна побачити з даних табл. 1, застосування алгоритмів стиснення є надзвичайно важливим інструментом як для ефективного зберігання аудіоінформації, так і для її передачі через канали зв’язку. Використання MPEG3-кодування дозволяє суттєво зменшити обсяг даних без критичної втрати якості, що робить цей підхід базовим стандартом у більшості сучасних мультимедійних систем.

Представимо у вигляді конкретних кроків алгоритм дослідження, який дозволяє визначати трансформанти перетворення Уолша-Адамара аудіоконтейнерів, що є найбільш вразливими до стиснення MPEG3. Застосування цього алгоритму дасть змогу ідентифікувати ті компоненти сигналу, які зазнають найбільших змін під час компресії, а отже, є критично важливими для забезпечення стійкості стеганографічних методів.

Крок 1. Ініціалізувати вектор

$$\begin{array}{r|cccc} i & 0 & 1 & \dots & N-1 \\ \hline \textit{Кількість} & 0 & 0 & \dots & 0 \end{array}. \qquad (6)$$

Крок 2. Обрати та зчитати аудіофайл у форматі без втрат FLAC. Здійснити стиснення обраного аудіофайлу із заданим бітрейтом і зберегти його у форматі з втратами MPEG-3.

Крок 3. Розбити обидві послідовності (без стиснення та із стисненням) на блоки розміру N. Знайти перетворення Уолша-Адамара згідно з (1) для блоків вихідної послідовності та блоків послідовності після стиснення алгоритмом MPEG-3.

Крок 4. Знайти різницю між трансформантами перетворення Уолша-Адамара блоків вихідної послідовності та послідовності із стиском.

Крок 5. Обрати трансфрманту перетворення Уолша-Адамара, яка зазнала найменших змін, інкрементувати значення кількості у (6), що відповідає номеру зазначеної трансформанти.

Крок 6. Якщо оброблено всі аудіофайли, зупин, інакше перейти до Кроку 2.

Крок 7. Здійснити представлення вектора (6) у вигляді стовпчикової діаграми.

Для наочності та кращого розуміння роботи запропонованого підходу алгоритм

дослідження стійкості трансформант перетворення Уолша-Адамара до стиснення MPEG3 буде представлено у вигляді блок-схеми (рис. 2). Такий графічний опис дозволяє чітко простежити послідовність етапів обробки аудіосигналу, від підготовки вхідних файлів і застосування перетворення до аналізу впливу стиснення на окремі трансформанти.

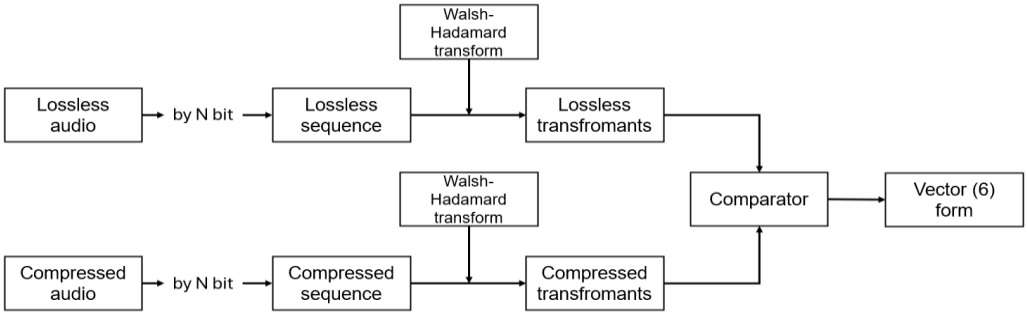


Рис. 2. Схема розробленого алгоритму пошуку трансформант перетворення Уолша-Адамара, що зазнають найбільшого спотворення через стиснення

Як видно з рис. 2, алгоритм обробки отримує два потоки аудіоданих: один у форматі FLAC, інший – у форматі MPEG3. Кожен з них розбивається на послідовності довжиною N елементів. Далі до кожної пари блоків застосовується перетворення Уолша-Адамара [12], основна мета якого – визначити позиції в послідовності, що найменше або найбільше змінюються під час атаки стисненням.

Індекси позицій, які зазнали найменших змін, інкрементуються у відповідному векторі (6), на основі якого будується діаграма частоти появи позицій у блоці, що найменше піддавалися змінам у процесі обробки всього набору даних. У цьому контексті менші значення на гістограмі вказують на статистично більшу стійкість позиції до змін під час стиснення, що дозволяє ідентифікувати трансформанти, найбільш придатні для вбудовування прихованої інформації.

На рис. 3 представлено графічне відображення результатів дослідження стійкості трансформант перетворення Уолша-Адамара до стиснення MPEG3.

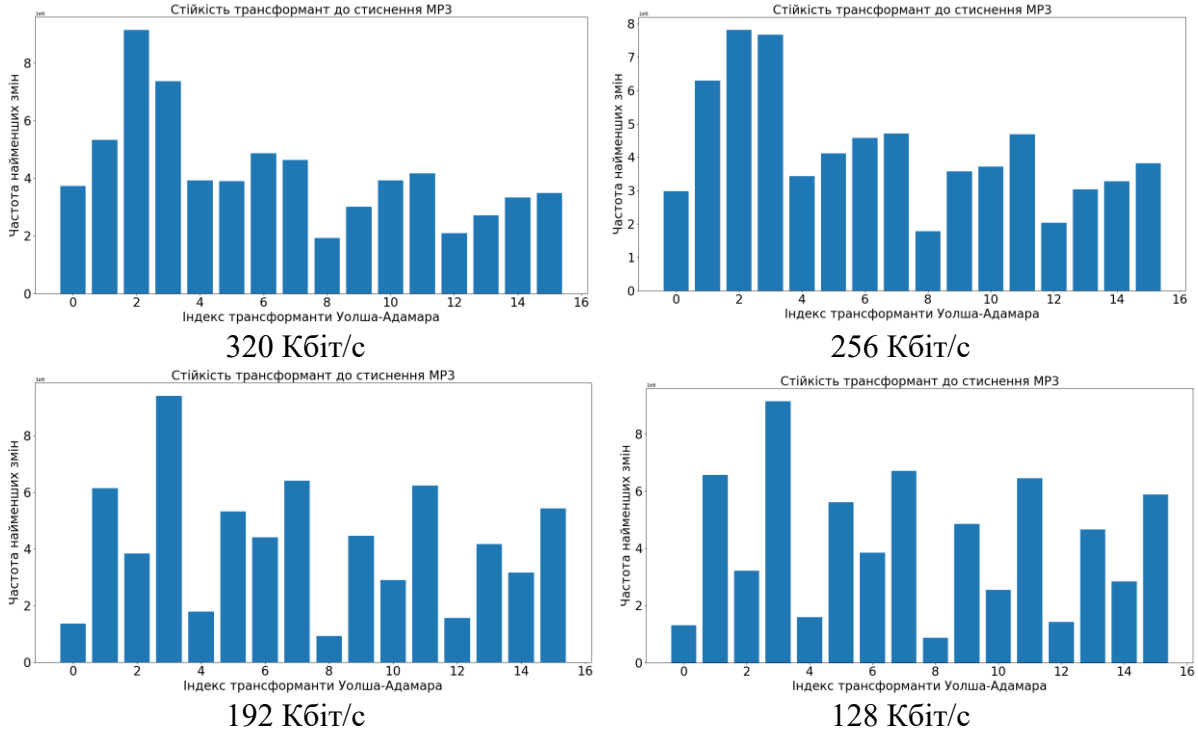


Рис. 3. Зміни трансформант перетворення Уолша-Адамара у аудіофайлах при стисненні MPEG3 для довжини блока $N = 16$

Аналіз представлених на рис. 3 гістограм показав закономірності зміни трансформант при стисненні MPEG3 з різними бітрейтами. Для бітрейтів 256 Кбіт/с і вище найменшому пошкодженню піддаються трансформанти з індексами 8 та 12, що робить їх пріоритетними для вбудовування додаткової інформації. При зниженні бітрейту до 192 та 128 Кбіт/с виявляється можливим також використовувати трансформанти 0 та 4 для вбудовування, проте навіть у цих умовах основною та найбільш стійкою залишається трансформанта 4. Таким чином, результати дослідження дозволяють виділити трансформанти, які забезпечують максимальну стійкість до втрат при стисненні та можуть бути рекомендовані для подальшого застосування методів кодового управління у аудіоконтейнерах. Відзначимо також, що дослідження змін трансформант перетворення Уолша-Адамара в аудіофайлах під впливом MPEG3-стиснення для більших довжин блока (наприклад, $N = 64$) є вкрай важливими. Збільшення розміру блока змінює спектральну роздільну здатність перетворення, по-іншому розподіляє енергію між трансформантами й може виявити інші, більш стійкі до компресії компоненти сигналу, що потенційно збільшує стійкість стеганографічного методу.

Для стислості ми наводимо на рис. 4 гістограму змін трансформант перетворення Уолша-Адамара, побудовану згідно до запропонованого нами алгоритму для бітрейту 128 Кбіт/с і довжини блока $N = 64$.

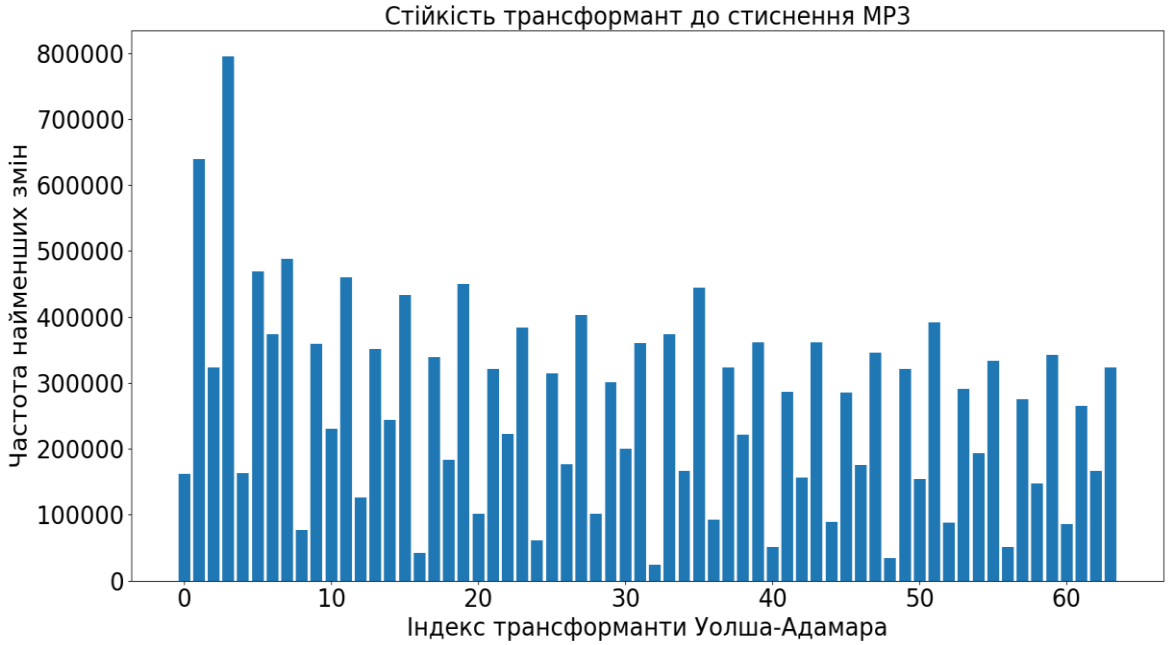


Рис. 4. Зміни трансформант перетворення Уолша-Адамара у аудіофайлах при стисненні MPEG3 для довжини блока $N = 64$

Аналіз представленої на рис. 4 гістограми для блоку довжиною 64 та бітрейту 128 кбіт/с показує, що найбільш стійкою до MPEG3-стиснення є трансформанта номер 32. Саме її рекомендується використовувати для вбудовування додаткової інформації. Водночас трансформанти 16, 24, 40, 48 та деякі інші також демонструють відносну стійкість і можуть слугувати альтернативними позиціями для вбудовування, проте вони поступаються трансформанті 32 за стабільністю та частотою найменших змін.

Висновки. Розв’язано важливу науково-прикладну задачу дослідження стійкості трансформант перетворення Уолша-Адамара в аудіоконтейнерах до атак стисненням на прикладі алгоритму MPEG3. Отримані результати дозволяють сформулювати такі висновки:

1. Проведено ґрунтовний аналіз сучасних методів аудіостеганографії, що показав наявність істотних недоліків існуючих рішень – низьку стійкість до атак

стисненням, значні обчислювальні витрати або обмежену універсальність. Це обґрунтувало доцільність дослідження нових підходів, зокрема методів на основі кодового управління.

2. Запропоновано та реалізовано алгоритм ідентифікації найбільш стійких до стиснення MPEG3 трансформант перетворення Уолша-Адамара. Алгоритм дозволяє статистично визначати ті компоненти спектру сигналу, які зазнають мінімальних змін під час компресії та, відповідно, є найбільш придатними для приховування інформації.

3. Експериментальні дослідження на базі широкого набору аудіофайлів у форматі HiFi довели, що трансформанти з індексами 8 та 12 стабільно демонструють найвищу стійкість до спотворень незалежно від бітрейту стиснення. Разом із тим, при нижчих бітрейтах (192 та 128 Кбіт/с) можливим є також використання трансформант 0 та 4, які зберігають відносну стійкість у цих умовах. Для блоків більшої довжини (64 елементи) найвищу стабільність до компресії показала трансформанта 32. Отримані результати підтверджують залежність оптимального вибору трансформант від параметрів стиснення та розміру блока.

4. Отримані результати підтверджують перспективність використання кодового управління в аудіостеганографії. Ідентифіковані стійкі трансформанти перетворення Уолша-Адамара можуть слугувати основою для побудови адаптивних алгоритмів приховування інформації, що поєднують високу стійкість до атак стисненням, невиявність та обчислювальну ефективність.

Таким чином, розроблений підхід формує науково обґрунтовану базу для подальшої розробки практично орієнтованих систем аудіостеганографії нового покоління, здатних забезпечувати надійний захист інформації навіть у середовищах із жорсткими обмеженнями на обчислювальні ресурси та активним застосуванням алгоритмів стиснення.

Список літератури

1. Babando A. K., Ahmad B. M. Data security using steganography. *LC International Journal of STEM*. 2022. Vol. 3, No. 4. P. 12-24.
2. Evsutin O., Melman A., Meshcheryakov R. Digital Steganography and Watermarking for Digital Images: A Review of Current Research Directions. *IEEE Access*. 2020. No. 8. P. 166589-166611. doi: 10.1109/ACCESS.2020.3022779
3. Кобозева А.А., Хорошко В.А. Анализ информационной безопасности. К.: ГУИКТ, 2010. 251 с.
4. Gopalan K. Audio steganography using bit modification. *International Conference on Multimedia and Expo. ICME'03. Proceedings. IEEE*. 2003. Vol. 1. P. I-629. DOI: 10.1109/icme.2003.1220996
5. Li J., Wang K., Jia X. A coverless audio steganography based on generative adversarial networks. *Electronics.* 2023. Vol. 12, No. 5. P. 1253. DOI 10.3390/electronics12051253
6. Wang J., Wang K. A novel audio steganography based on the segmentation of the foreground and background of audio. *Computers and Electrical Engineering*. 2025. Vol. 123. P. 110026. DOI 10.1016/j.compeleceng.2024.110026
7. Su W. et al. Efficient audio steganography using generalized audio intrinsic energy with micro-amplitude modification suppression. *IEEE Transactions on Information Forensics and Security*. 2024. Vol. 19. P. 6559-6572. DOI: 10.1109/tifs.2024.3417268
8. Feng Y. et al. A robust coverless audio steganography based on differential privacy clustering. *IEEE Transactions on Multimedia*. 2025. DOI: 10.1109/tmm.2025.3543107
9. Kobozeva A.A., Sokolov A.V. Robust Steganographic Method with Code-Controlled Information Embedding. *Problemele energeticii regionale*. 2021. No. 4 (52). P. 115-130. DOI: 10.52254/1857-0070.2021.4-52.11
10. Kobozeva A.A., Sokolov A.V. Theoretical foundations for constructing effective codewords for the code-controlled information embedding steganographic method. *Radiotekhnika.* 2021. 4(207). P. 27–39. DOI: 10.30837/rt.2021.4.207.02.

11. Kobozeva A.A., Sokolov A.V. Steganographic Method with Code Control of Information Embedding Based on Multi-level Code Words. *Radioelectronics and Communications Systems*. 2023.Vol. 66, No. 4. P. 173-189. DOI: 10.3103/s0735272723040052
12. Sokolov A.V., Ihnatenko O.O., Balandina N.M. Increasing the Efficiency of Blind Decoding of the Steganographic Method with Code Control of Additional Information Embedding. Problems of regional energetics. 2024. Vol. 62, No. 2. P. 121-137. DOI: 10.52254/1857-0070.2024.2-62.11
13. Sokolov A.V. Multiple access steganographic method based on code control and frequency arrangements. *Informatics and Mathematical Methods in Simulation*. 2021. Vol. 11, No. 3. P. 147-161. DOI: 10.15276/imms.v11.no3.147
14. Ahmed N., Rao K. R. Walsh-Hadamard transform. Orthogonal transforms for digital signal processing. Berlin, Heidelberg : Springer Berlin Heidelberg, 1975. P. 99-152.
15. Bosi M. et al. ISO/IEC MPEG-2 advanced audio coding. *Journal of the Audio engineering society*. 1997. Vol. 45, No. 10. P. 789-814.
16. Kosaka A. et al. Design of Ogg Vorbis decoder system for embedded platform. *IEICE Transactions on Fundamentals*. 2005. Vol. 88, No. 8. P. 2124-2130. DOI: 10.1093/ietfec/e88-a.8.2124
17. Shlien S. Guide to MPEG-1 audio standard. *IEEE Transactions on Broadcasting*. 2002. Vol. 40, No. 4. P. 206-218. DOI: 10.1109/11.362938
18. Wang Y., Vilermo M. Modified discrete cosine transform: Its implications for audio coding and error concealment. *Journal of the Audio Engineering Society.* 2003. Vol. 51, No. 1/2. P. 52-61.

RESEARCH OF THE ROBUSTNESS OF WALSH-HADAMARD TRANSFORMANTS AGAINST MPEG3 COMPRESSION FOR AUDIO STEGANOGRAPHY

[1]A.V. Sokolov, [2]M.V. Khymenko

[1]National University "Odesa Law Academy"
23, Fontanska doroha, Odesa, 65009, Ukraine
[2]National Odesa Polytechnic University
1, Shevchenko Ave., Odesa, 65044, Ukraine

The paper presents a new direction in the development of audio steganography – adaptation and justification of code control methods to increase the resistance of covert messages in audio containers to modern compression attacks, in particular to MPEG3. The authors perform a systematic analysis of modern approaches to audio steganography and emphasize their limitations – loss of covert data during compression or excessive demands on computing resources, which complicates their use in mobile and embedded systems. In this context, it is proposed to adapt the concept of code control to audio containers, which has already demonstrated high effectiveness for images. This approach opens up the possibility of controlled data embedding into stable signal components, combining stealth, resistance to attacks and computational efficiency. The key contribution of the paper is the development of an algorithm for identifying Walsh-Hadamard transformants, which are minimally distorted during MPEG3 compression. The algorithm sequentially compares pairs of blocks of the original (FLAC) and compressed (MPEG3) signals, calculates the Walsh-Hadamard transform, accumulates statistics of changes and visualizes the frequencies of minimum distortion of the Walsh-Hadamard transformants in the form of histograms, which allows to objectively highlight priority positions for embedding. The experiments were performed on a wide set of HiFi audio recordings with testing of several bitrates (320, 256, 192, 128 kbit/s) and different block lengths, which ensures the representativeness and reliability of the conclusions. The experimental results demonstrate a stable pattern: transformants with indices 8 and 12 consistently exhibit the highest resistance to distortion during compression and are prioritized for embedding hidden data; when the bitrate is reduced to 192 and 128 kbit/s, in addition to the above transformants, it is practically permissible (and advisable in a number of scenarios) to use transformants 0 and 4, which expands the set of safe positions for code control. In addition, when the block length is increased to 64 elements, it is shown that transformant 32 is stable. The practical significance of the research is that the selected stable transformants can become the basis for creating adaptive, highly effective steganography algorithms with code control: they allow combining invisibility, resistance to common compression algorithms (including MPEG3) and low computational costs, which makes the proposed approach suitable for implementation in real systems. The paper also opens up prospects for further research – expanding the analysis to other compression algorithms and multimedia formats, integrating machine learning methods for automatic adaptation of codewords and developing new steganographic schemes.

Keywords: audio steganography, code control, Walsh-Hadamard transform, steganographic message, compression robustness, MPEG3 compression, transformants, information security.

DOI 10.15276/imms.v15.no4.636
УДК 004.75:004.77

Informatics and Mathematical Methods in Simulation
Vol.15 (2025), No. 4, pp. 636-646

ПОРІВНЯЛЬНИЙ АНАЛІЗ БЕЗПЕКОВИХ ПАТЕРНІВ ХМАРНИХ ПЛАТФОРМ У КОНТЕКСТІ НАУКОВОЇ ВІДТВОРЮВАНОСТІ

О. Г. Трофименко[1], П. О. Чикунов[1], Д. Ю. Астахов[1], Ю. В. Молоканов, Т. А. Фаріонова[2]

[1]Національний університет «Одеська юридична академія»
23, Фонтанська дорога, Одеса, 65009, Україна
[2]Національний університет кораблебудування імені адм. Макарова
9, Героїв України пр., Миколаїв, 54007, Україна
Emails: trofymenko@onua.edu.ua, pavel@onua.edu.ua, astakhovnil@gmail.com, molokanov9@gmail.com, tetyana.farionova@nuos.edu.ua

Дослідження присвячене системному аналізу реалізації ключових безпекових патернів у хмарних платформах Amazon Web Services, Microsoft Azure та Google Cloud Platform з урахуванням їх придатності для наукових досліджень у галузі комп'ютерних наук. Актуальність дослідження зумовлена потребою в забезпеченні безпеки, ізоляції, автоматизації та відтворюваності інфраструктури в умовах мультихмарного середовища. Мета дослідження полягає у систематизації та порівняльному аналізі ключових безпекових патернів у хмарних платформах AWS, Azure та GCP з урахуванням їх ефективності, обмежень та придатності для забезпечення відтворюваності, ізоляції та етичної відповідності в наукових дослідницьких середовищах. Методологія дослідження базується на системному підході, який включає порівняння конфігурацій хмарних платформ на основі практичних кейсів, застосування інструментів статичного аналізу (Checkov, tfsec) для оцінки безпеки IaC-коду, моделювання поведінки системи у сценаріях відмови та атак для емпіричної перевірки стійкості інфраструктури. У статті розглянуто три практичні кейси: 1) для AWS – реалізація Role Vending Machine для автоматизованого призначення тимчасових ролей з обмеженими правами, інтеграція з GitHub Actions, Checkov та IAM Access Analyzer; 2) для Azure – багатозонна архітектура з PIM, GitOps, ARM-шаблонами та мережевою сегментацією для моделювання відмовостійкості; 3) для GCP – використання Shared VPC, ізольованих проєктів, TPU та Vertex AI Experiments для задач NLP і глибокого навчання. Порівняльний аналіз показав: принцип найменших привілеїв значно знижує ризики витоку прав; відсутність логування та аудиту створює "мертві" права, які можуть бути точками входу для атак; ізоляція середовищ (через окремі акаунти, підписки, проєкти) критично важлива для запобігання перехресному впливу; управління змінами через CI/CD або GitOps забезпечує повторюваність і контроль; Threat modeling дозволяє виявити уразливості, пов'язані з розміщенням ресурсів; регуляторні вимоги (GDPR, шифрування, геозони) мають бути враховані на етапі архітектурного планування. Проведене дослідження дозволяє зробити висновок, що хмарна інфраструктура, налаштована з урахуванням принципів безпеки, автоматизації та етичної відповідності, може стати повноцінним науковим інструментом, а не лише технічним середовищем. Водночас нехтування цими принципами створює ризики компрометації даних, порушення регуляторних норм та втрати достовірності результатів. Наукова новизна роботи полягає у порівняльному аналізі безпекових моделей трьох хмарних платформ у контексті забезпечення достовірності та етичної відповідності наукових експериментів. Практична значущість результатів полягає у можливості їх використання для побудови безпечної, масштабованої та відтворюваної хмарної інфраструктури в наукових проєктах, у професійній діяльності DevOps-команд, які працюють з мультихмарними середовищами, як основа для розробки етичних та регуляторно відповідних архітектур у сфері комп'ютерних наук.

Ключові слова: хмарні обчислення, мультихмарне середовище, інфраструктура як код, DevOps, DevSecOps, безпекові патерни, CI/CD, GitOps, хмарна архітектура.

Вступ. Хмарні обчислювальні платформи стали фундаментальним компонентом сучасної наукової інфраструктури, особливо в галузі комп'ютерних наук. Вони

забезпечують масштабоване середовище для реалізації алгоритмічних моделей, високопродуктивних симуляцій, розгортання нейронних мереж, зберігання великих обсягів даних та автоматизації експериментальних циклів. В умовах обмежених локальних ресурсів і потреби в гнучкому масштабуванні платформи Amazon Web Services (AWS), Microsoft Azure та Google Cloud Platform (GCP) пропонують інструменти, які дозволяють не лише ефективно керувати обчисленнями, а й забезпечити наукову точність, повторюваність та відтворюваність результатів.

У контексті стрімкого розвитку хмарних технологій та DevOps-практик, хмарна інфраструктура стала критично важливою для реалізації високопродуктивних обчислень, зберігання даних і автоматизації наукових експериментів.

Аналіз досліджень та публікацій показує активне зростання інтересу до застосування мультихмарних архітектур, реалізації ключових безпекових патернів у хмарних платформах. У статті [1] виконано огляд концепції «Інфраструктура як код» (Infrastructure as Code, IaC) для мультихмарних налаштувань, зосереджуючись на модульній архітектурі, стандартизації інструментів, управлінні, інтеграції безпеки та автоматизації через конвеєри CI/CD (Continuous Integration / Continuous Delivery). Дослідження [2] підкреслює трансформаційний потенціал використання кількох хмарних архітектур у сучасних корпоративних середовищах, наголошуючи на їхній ролі в досягненні бізнес-гнучкості, оптимізації витрат та операційної ефективності. У статтях [3, 4] обговорюються найкращі практики впровадження автоматизації безпеки за допомогою практик статичного та динамічного тестування безпеки застосунків (SAST/DAST), автоматизовані перевірки відповідності, захист під час виконання за допомогою сканера безпеки інфраструктури як коду, рішення для безпеки контейнерів та системи виявлення аномалій на основі поведінки. Дослідження [5] заглиблюється в перетин IaC, CI/CD та оркестрації Kubernetes у мультихмарних DevOps-конвеєрах. Робота [6] аналізує еволюцію мультихмарної стратегії і роль Terraform у хмарній оркестрації з невеликим порівнянням з іншими інструментами хмарної оркестрації. У статті [7] розглядається концепція інфраструктури як коду з позиції її реалізації за допомогою Terraform, висвітлюються переваги та проблеми IaC у сучасних хмарних середовищах. У роботі [8] класифіковано широко визнані методи безпеки Terraform, поширені в галузі для популярних хмарних провайдерів, таких як AWS, Azure та Google Cloud.

Проведений аналіз засвідчує наявність ґрунтовної дослідницької бази щодо хмарних платформ. Водночас, попри широке застосування хмарних сервісів у наявних публікаціях недостатньо висвітлено питання інфраструктурної безпеки, ізоляції середовищ, управління ролями та відтворюваності конфігурацій у контексті наукової методології. Зокрема, відсутній системний аналіз того, як принцип найменших привілеїв, аудит доступу та моделювання загроз впливають на достовірність і безпечність експериментальних результатів.

Метою даної роботи є систематизація та порівняльний аналіз реалізації ключових безпекових патернів у хмарних платформах Amazon Web Services, Microsoft Azure та Google Cloud Platform з урахуванням їх ефективності, обмежень та придатності для забезпечення відтворюваності, ізоляції та етичної відповідності в наукових дослідницьких середовищах.

Для досягнення поставленої мети визначено такі завдання:

– проаналізувати сучасні наукові публікації та дослідження для виявлення ключових напрямів, тенденцій та проблем у застосуванні патернів хмарної інфраструктури для наукових досліджень у галузі комп’ютерних наук, з акцентом на безпеку, ізоляцію, автоматизацію та відповідність етичним нормам;

– дослідити переваги інфраструктури як коду для DevOps;

– виконати порівняльний аналіз трьох провідних хмарних платформ AWS, Azure, GCP щодо їх придатності для дослідницьких задач;

– продемонструвати, як правильно налаштовані хмарні середовища можуть забезпечити відтворюваність, масштабованість та захист даних, необхідні для наукової достовірності.

Переваги інфраструктури як коду для DevOps. Розвиток хмарних технологій є одним із ключових напрямів сучасної інформатизації, що визначає ефективність цифрової трансформації бізнесу, науки та державного управління. Останнє десятиріччя позначене переходом від простих моделей оренди обчислювальних ресурсів до комплексних платформ, орієнтованих на масштабованість, автоматизацію та інтеграцію із сучасними методологіями розробки. Якщо на початкових етапах розвитку хмарні обчислення концентрувалися на забезпеченні гнучкого доступу до інфраструктури у форматі IaaS (Infrastructure as a Service), то наразі спостерігається зростання популярності платформних рішень (PaaS, Platform as a Service) та сервісів, орієнтованих на мікросервісну архітектуру, контейнеризацію й оркестрацію за допомогою системи Kubernetes.

Важливим трендом є поширення мультихмарних та гібридних стратегій, які дозволяють поєднувати ресурси різних провайдерів з метою оптимізації витрат, підвищення відмовостійкості та уникнення ризику «vendor lock-in». Це вимагає розробки нових інструментів управління, здатних забезпечити уніфікований підхід до оркестрації інфраструктури, незалежно від її розподілу між AWS, Microsoft Azure, Google Cloud чи приватними дата-центрами [6]. Одночасно з цим посилюється акцент на безпеці, дотриманні нормативних вимог та автоматизації процесів відповідно до DevOps-практик [7]. Як результат формується потреба у системному підході до управління інфраструктурою, що дозволяє скоротити час розгортання середовищ, підвищити їх передбачуваність та якість обслуговування.

У цьому контексті концепція «Інфраструктура як код» стала однією з базових технологічних передумов переходу до інженерії сучасних хмарних середовищ. IaC описує інфраструктуру у вигляді декларативних чи імперативних конфігураційних файлів, що дозволяє автоматизувати процеси створення, масштабування, модифікації та видалення ресурсів [8]. Тим самим IaC дозволяє командам автоматизувати та керувати інфраструктурою за допомогою програмного коду, забезпечуючи узгодженість, масштабованість та швидше розгортання. Такий підхід забезпечує прозорість і відтворюваність усіх дій, оскільки інфраструктура визначається так само, як і програмний код, а її стан зберігається у системах контролю версій.

На рис. 1 зображено типовий робочий процес IaC, який охоплює основні компоненти та взаємозв'язки між ними.

Рис. 1. Інфраструктура як код

Процес починається з дій розробника або інженера DevOps, який створює код інфраструктури, використовуючи спеціалізовані мови опису, як-от: Terraform, Ansible або CloudFormation. Цей код містить декларативні інструкції щодо створення, налаштування та управління обчислювальними ресурсами, включаючи сервери, мережі, бази даних та інші компоненти. Після написання, код передається до системи контролю

версій, найчастіше Git, що дозволяє зберігати історію змін, здійснювати пул-реквести та забезпечувати колективну роботу над інфраструктурними сценаріями.

Затверджений код надходить до автоматизаційного сервера або API-рушія, який виконує його інтерпретацію та застосування. Цей компонент відповідає за інтеграцію з хмарними платформами або локальними дата-центрами, забезпечує розгортання інфраструктури відповідно до заданих параметрів. У результаті інфраструктура може бути реалізована як у хмарному середовищі, так і на локальних серверах, що забезпечує гнучкість та масштабованість IT-рішень.

Важливою особливістю IaC є її здатність забезпечувати ідентичність середовищ, зменшувати ризики людських помилок та сприяти впровадженню практик DevOps, таких як CI/CD. Тим самим інфографіка демонструє не лише технічну послідовність дій, а й концептуальну модель переходу від ручного управління інфраструктурою до її повної автоматизації через код.

Використання IaC у рамках DevOps дає змогу досягати низки важливих переваг. По-перше, автоматизація розгортання суттєво зменшує кількість людських помилок, характерних для ручного налаштування інфраструктури [9]. По-друге, забезпечується повторюваність середовищ, що критично важливо для тестування, розробки та продуктивної експлуатації [10]. По-третє, IaC інтегрується з CI/CD-процесами, дозволяючи створювати інфраструктуру «на вимогу» у відповідь на зміни в програмному забезпеченні. По-четверте, завдяки централізованому управлінню станом та конфігураціями підвищується рівень безпеки й керованості мультихмарних середовищ [11].

Крім того, IaC дозволяє організаціям оптимізувати витрати на хмарні ресурси, програмно керуючи ресурсами, забезпечуючи автоматичне вилучення невикористаних ресурсів. Це запобігає непотрібним витратам на хмарні ресурси та їх втратам. Сучасні інструменти IaC підтримують мультихмарні середовища, дозволяючи організаціям безперешкодно розгортати робочі навантаження в AWS, Azure та Google Cloud. Це дозволяє уникнути прив'язки до постачальника та підвищує гнучкість [9].

Отже, IaC є важливою практикою в сучасному DevOps, яка пропонує швидкість, масштабованість, автоматизацію та узгодженість в управлінні інфраструктурою. Використовуючи інструменти IaC, як-от: Terraform, Ansible та AWS CloudFormation, DevOps-фахівці можуть оптимізувати операції, знизити витрати та покращити доставку програмного забезпечення.

Порівняння ключових безпекових підходів та практик AWS, Azure та GCP. У кожній із найпотужніших хмарних платформ – AWS, Azure і GCP – реалізовано унікальні механізми, які дозволяють адаптувати інфраструктуру до специфіки дослідницьких задач.

Terraform як інструмент IaC дозволяє декларативно описувати, створювати та керувати інфраструктурою в хмарних середовищах. Його головна перевага – уніфікований підхід до роботи з різними хмарами, зокрема AWS, Azure та GCP.

Terraform використовує спеціальні провайдери для кожної хмарної платформи: aws – для Amazon Web Services, azurerm – для Microsoft Azure, google – для Google Cloud Platform. Користувач описує бажаний стан інфраструктури у .tf файлах, а Terraform самостійно визначає, які ресурси створити, змінити або видалити. При цьому один і той самий інструмент і синтаксис дозволяє працювати з різними хмарами (табл. 1), що спрощує гібридні або мультихмарні архітектури.

У табл. 2 наведено узагальнені ключові безпекові підходи та практики у розглянутих трьох провідних хмарних платформ у розрізі категорій, що мають вирішальне значення для побудови безпечної та відтворюваної інфраструктури за допомогою Terraform у мультихмарному середовищі. Порівняння дозволяє виявити як спільні стратегічні підходи, так і відмінності в реалізації конкретних механізмів, що мають критичне значення в наукових та інженерних обчисленнях.

Таблиця 1.

Особливості інтеграції з кожною платформою

Платформа	Провайдер	Приклади ресурсів	Особливості
AWS	aws	aws_instance, aws_s3_bucket, aws_lambda_function	Найбільш зрілий провайдер, підтримує майже всі сервіси AWS
Azure	azurerm	azurerm_virtual_machine, azurerm_storage_account, azurerm_function_app	Потребує автентифікацію через Azure CLI або Service Principal
GCP	google	google_compute_instance, google_storage_bucket, google_cloudfunctions_function	Підтримка через Google Cloud SDK або ключі доступу

Таблиця 2.

Порівняльна таблиця безпекових патернів AWS, Azure та GCP

Категорія	AWS	Asure	GCP
Найменші привілеї	IAM Role Vending Machine (RVM)	Privileged Identity Management (PIM)	Custom IAM Policies + Time-bound Access
Логування та аудит	IAM Access Analyzer, CloudTrail	Activity Logs, PIM Audit	VPC Flow Logs, Cloud Audit Logs
Ізоляція середовищ	Окремі VPC, окремі акаунти	Окремі підписки, групи ресурсів	Shared VPC + окремі проєкти
Управління змінами	CI/CD з GitHub Actions + Checkov	GitOps + ARM Templates	Terraform + Cloud Build
Моделювання загроз	Placement-aware EC2, network ACLs	WAF, NSG, segmentation	Co-location analysis, TPU isolation
Регуляторна відповідність	KMS, region control, data lifecycle	Encryption at rest/in transit, GDPR zones	Data residency, DLP API
Економічна ефективність	Auto-stop EC2, spot instances	Budget alerts, cost analysis	Preemptible VMs, region-aware planning

У категорії найменших привілеїв AWS демонструє високий рівень автоматизації, завдяки реалізації патерна Role Vending Machine (RVM), який забезпечує контрольований процес призначення тимчасових ролей з обмеженими правами доступу в хмарній інфраструктурі. Основною перевагою цього підходу є зменшення ризику надмірного доступу або помилок прав, а також спрощення аудиту через стандартні інструменти (IAM Access Analyzer, Checkov). IAM (Identity and Access Management) є системою керування ідентифікацією та доступом у хмарній платформі. Важливо розуміти і зважати на виклики, пов’язані зі складністю підтримки шаблонів і рев’ю-процесів при великій кількості проєктів. Azure натомість пропонує Privileged Identity Management (PIM) як сервіс із вбудованою тимчасовою ескалацією прав доступу, що дозволяє обмежити постійне використання адміністративних ролей. Тим самим Azure через PIM надає гнучке управління привілейованими правами, дозволяє активувати їх лише на час, необхідний для адміністративних дій. Це корисно, коли дослідницька команда має змінні ролі, і потрібно обмежити загрозу постійного адміністративного доступу. Але недоліком є те, що PIM залежить від правильного налаштування політик та умов активації; якщо активація занадто проста або не перевіряється, роль може бути зловживана. GCP реалізує подібну концепцію через кастомізовані IAM-політики з можливістю тимчасового надання доступу, часто інтегровані з TTL-механізмами, наприклад через Cloud Scheduler або сторонні інтеграції. Це дозволяє зберігати автономію обчислювальних середовищ, сприяє повторюваності експериментів, бо всі команди працюють в узгоджених мережевих умовах. Проте виклик полягає в тому, що не всі сервіси GCP повністю підтримують Shared VPC або мають обмеження щодо сервісних агентів, підмереж або прав, що можуть бути делегованими. Також, при

масштабуванні мережевих правил і firewall конфігурацій витрати на адміністрування зростають, і ризик помилкових правил або надмірних дозволів може збільшуватися.

У сфері логування та аудиту всі три платформи надають інструменти для детального моніторингу змін і дій користувачів, але з відмінностями в інтеграції та гнучкості. AWS поєднує IAM Access Analyzer з CloudTrail для детекції надмірних прав та історії доступу. Azure забезпечує аудит через Activity Logs і PIM Audit, з фокусом на привілейовані дії. GCP пропонує Cloud Audit Logs разом із VPC Flow Logs, що робить акцент як на дії користувачів, так і на мережевих з'єднаннях.

Ізоляція середовищ реалізується через архітектурні механізми розмежування: AWS використовує окремі облікові записи та VPC, Azure – підписки й групи ресурсів, тоді як GCP робить ставку на поєднання окремих проєктів із Shared VPC для централізованого управління мережею. Кожен із підходів має свої переваги в масштабованості та адмініструванні – зокрема, підхід GCP дозволяє зменшити дублювання мережевої конфігурації без втрати ізоляції.

Щодо управління змінами, то всі три платформи підтримують інфраструктурний підхід з використанням Terraform, однак мають різні точки інтеграції з CI/CD. AWS застосовує GitHub Actions у поєднанні з інструментами статичного аналізу (наприклад, Checkov), що дозволяє виявляти потенційні проблеми в конфігураціях ще до їхнього застосування. Azure традиційно використовує ARM-шаблони та GitOps-підходи з прив'язкою до Azure DevOps. GCP орієнтується на поєднання Terraform з Cloud Build, що забезпечує нативну інтеграцію з іншими сервісами GCP у пайплайні.

У контексті моделювання загроз кожна платформа реалізує власні механізми попередження ризиків: AWS враховує політику розміщення ресурсів у хостах (placement), мережеві ACL і фізичну ізоляцію в EC2; Azure фокусується на використанні WAF, мережевих груп безпеки (NSG) і сегментації, а GCP надає розширені можливості аналізу розміщення процесів на спільних вузлах (особливо для TPU) і підтримку зонованої ізоляції. Це набуває особливої важливості у дослідницьких сценаріях, де є ризик атак побічними каналами в мульти-тенантних середовищах.

У частині регуляторної відповідності всі три платформи підтримують вимоги до захисту даних як у спокої, так і в русі. AWS забезпечує контроль регіонів зберігання та гнучке управління життєвим циклом даних у поєднанні з KMS. Azure надає механізми шифрування та підтримку збереження даних у зонах, що відповідають вимогам GDPR. GCP дозволяє реалізувати політику data residency та надає інструменти на кшталт DLP API для захисту конфіденційної інформації.

Економічна ефективність управління інфраструктурою також варіюється залежно від платформи. AWS пропонує автозупинку EC2 і використання spot-інстансів для зниження витрат. Azure реалізує моніторинг бюджету та аналітику витрат як частину Azure Cost Management. GCP, своєю чергою, оптимізує витрати за рахунок preemptible VM і планування розміщення ресурсів у найвигідніших регіонах.

Тим самим аналіз показує, що кожна з хмарних платформ має власні сильні сторони та підходи до вирішення спільних викликів безпеки, що формують ґрунт для застосування Terraform у мультихмарному науковому середовищі. Обґрунтований вибір залежить від вимог до ізоляції, регуляторної відповідності, масштабу експериментів і операційної стійкості.

Порівняння продуктивності та безпеки при реалізації основних безпекових патернів у хмарних середовищах AWS, Azure та GCP. Для проведення порівняльного аналізу було вибрано експериментальний підхід із побудовою типових сценаріїв розгортання інфраструктури у кожному з трьох хмарних середовищ – AWS, Azure та GCP – із використанням однакової логіки розгортання за допомогою Terraform. Усі експерименти виконувались із метою верифікації функціональності, вимірювання операційної продуктивності, а також оцінки рівня безпеки, реалізованого через відповідні нативні сервіси.

Застосування уніфікованих сценаріїв і автоматизованих засобів розгортання забезпечує наукову відтворюваність експериментів, оскільки створює можливість їх повторення в аналогічних умовах іншими дослідниками.

Кожна реалізація мала такі базові компоненти: мережеву інфраструктуру з ізоляцією середовищ; обчислювальні ресурси (2 віртуальні машини) для запуску навантаження; політики доступу з реалізацією найменших привілеїв; логування та моніторинг змін; пайплайни змін (CI/CD); активацію тимчасових привілеїв адміністратора. Розгортання інфраструктури проводилось через модульну архітектуру Terraform, з використанням однакових патернів розгортання (адаптованих під кожного провайдера) із відкритих репозиторіїв GitHub та офіційних модулів з Terraform Registry. Наприклад, модуль terraform-aws-vpc для AWS, Azure Network Module для Azure та terraform-google-network для GCP.

Для емпіричної оцінки продуктивності було застосовано таку стратегію:

– час повного розгортання інфраструктури (команда terraform apply) вимірювався для кожної платформи з однаковою конфігурацією (2 віртуальні машини, окрема мережа, IAM/role/policy); в середньому: AWS ~95 с, Azure ~123 с, GCP ~102 с;

– продуктивність середовища виконання вимірювалась при запусканні однакових бенчмарків CPU/GPU (Sysbench, MLPerf inference) на екземплярах із однаковими характеристиками (4 vCPU, 16 ГБ RAM, Linux Ubuntu 22.04 LTS). У CPU-бенчмарках AWS c5.large та GCP n2-standard-4 показали близьку продуктивність (~26k events/sec), тоді як Azure D4s_v3 відставала (~23k events/sec). У GPU-тестах AWS p3 і GCP A100 продемонстрували схожу продуктивність у TensorFlow-бенчмарках, проте GCP показала меншу латентність за рахунок TPU;

– для латентності політик доступу вимірювалась затримка між запитом тимчасових прав (через RVM, PIM, або IAM TTL) та фактичним набуттям прав доступу: AWS (RVM + OIDC) ~15-20 с, Azure (PIM) ~45-60 с, GCP (IAM TTL via Cloud Scheduler) ~25-30 с (рис. 2).

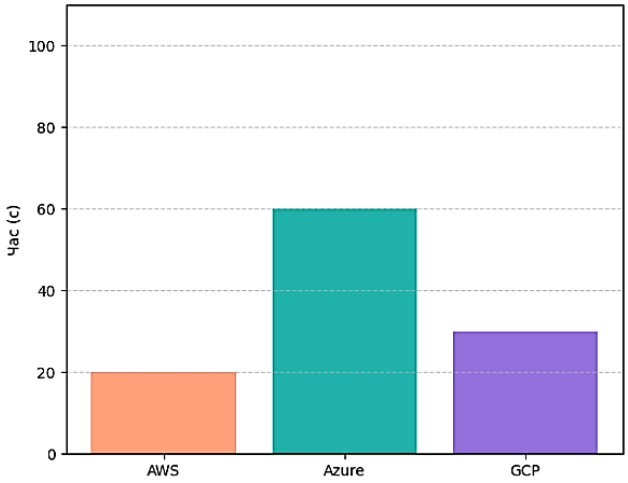


Рис. 2. Затримка надання тимчасових привілеїв доступу в хмарних провайдерах

Наведена на рис. 2 діаграма демонструє середню латентність (затримку) між запитом тимчасових прав доступу (через механізми RVM, PIM або TTL-based IAM) та моментом фактичного надання прав. Найшвидший час надання доступу продемонструвала AWS (приблизно 20 с) завдяки механізму Role Vending Machine з інтеграцією OIDC. Найбільшу затримку зафіксовано у Microsoft Azure (понад 60 с) через архітектурні особливості PIM, що може негативно впливати на гнучкість доступу в DevSecOps-сценаріях.

Для порівняння рівня безпеки було використано два підходи: статичний аналіз конфігурацій і перевірка поведінки системи у сценаріях відмови та атаки. Статичний аналіз здійснювався за допомогою Checkov та tfsec, зокрема для перевірки: відсутності

надмірних дозволів в IAM політиках, обов'язковості шифрування дисків і трафіка, увімкнення логування (CloudTrail / Audit Logs / Activity Logs). Другий підхід до оцінювання рівня безпеки хмарної інфраструктури стосувався емпіричного тестування поведінки системи у сценаріях відмови та потенційних атак [12]. На відміну від статичного аналізу конфігурацій, цей метод орієнтований на динамічну перевірку реакції середовища на порушення безпеки, що дозволяє виявити неочевидні вразливості, пов’язані з логікою доступу, обробкою помилок та збереженням аудиту [13]. Зокрема, було реалізовано низку контрольованих симуляцій, які моделюють типові загрози:

– спроба доступу до ресурсу без належних прав дозволяє перевірити ефективність механізмів автентифікації та авторизації, а також виявити можливі обхідні шляхи до критичних компонентів;

– зміна конфігурації мережі або параметрів міжмережевого екрану (firewall) імітує внутрішню або зовнішню атаку, спрямовану на порушення ізоляції середовища, що дає змогу оцінити стійкість системи до несанкціонованих змін;

– видалення або модифікація журналів аудиту (наприклад, CloudTrail, Audit Logs, Activity Logs) дозволяє визначити, чи здатна система виявити спроби приховати сліди активності, що критично важливо для забезпечення цілісності та трасування подій.

Застосування цього підходу дозволяє оцінити ефективність реалізованих патернів у реальних умовах та виявити потенційні прогалини, які не фіксуються засобами статичного аналізу. Він є важливим елементом комплексного аудиту безпеки хмарних рішень, особливо в контексті багатоплатформного середовища, де взаємодія компонентів може створювати складні сценарії ризику.

Після впровадження всіх патернів: в AWS блокування доступу спрацьовувало миттєво, логування відбувалось через CloudTrail, IAM Access Analyzer фіксував спроби зміни політик; в Azure аналогічну функцію виконував PIM Audit + Activity Log; в GCP – Cloud Audit Logs + Policy Analyzer. У результаті статичного аналізу найменше порушень виявлено в AWS (середній security score – 94/100), GCP – 91/100, Azure – 88/100 (через складніші ARM шаблони, які складно перевіряти автоматично) (рис. 3).

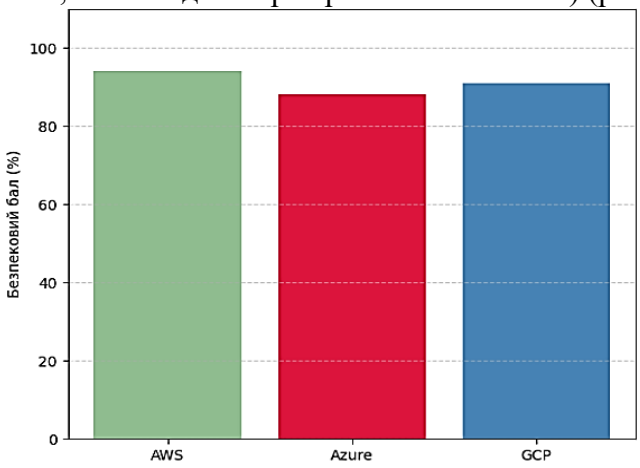


Рис. 3. Оцінка безпеки конфігурації хмарної інфраструктури

Статичний аналіз безпеки конфігурацій, розгорнутих у трьох провайдерах, із використанням інструментів Checkov та tfsec, охоплював політики доступу, наявність шифрування, активацію логування та відповідність рекомендаціям безпекових фреймворків (CIS Benchmarks). AWS продемонстрував найвищу оцінку (94%) завдяки чіткій структурі IAM та розвиненим сервісам моніторингу. Найменша оцінка у Azure (88%) зумовлена переважно складністю аналізу ARM-шаблонів та частковою відсутністю автоматизованого логування змін у політиках.

Результати та обговорення. Конкретні приклади налаштування інфраструктури за допомогою Terraform у трьох провідних хмарних платформах AWS, Azure та GCP продемонстрували різні підходи до безпеки, розмежування доступу, ізоляції середовищ

і централізованого управління мережею, зокрема через механізми RVM, PIM і Shared VPC відповідно. У випадку з AWS приклад реалізації Role Vending Machine довів можливість централізованого та автоматизованого надання прав доступу через опис ролей у вигляді коду, що в поєднанні з перевірками безпеки (IAM Access Analyzer, Checkov) забезпечує як гнучкість експериментального середовища, так і захист від неконтрольованого розширення прав. Аналіз Azure продемонстрував важливість привілейованого управління доступом у наукових системах із підвищеним рівнем відповідальності. Застосування Azure PIM у поєднанні з ARM-шаблонами забезпечило не лише ізоляцію середовищ, а й контроль над тимчасовими правами адміністративного рівня, що є ключовим для захисту життєвого циклу дослідження. У Google Cloud Platform приклад Shared VPC дав змогу реалізувати централізовану мережеву архітектуру з різними обчислювальними доменами, що дозволило виконувати порівняльні експерименти з мовними моделями в узгоджених умовах, забезпечуючи реплікованість результатів і контроль над витратами.

Поєднання різних патернів у мультихмарному середовищі дає змогу досягти високої безпеки, ізоляції, відтворюваності та контролю над витратами, якщо реалізовано належним чином. Водночас саме грамотне налаштування, з урахуванням принципів безпеки, автоматизації та відтворюваності, перетворює хмару на повноцінний науковий інструмент. Нехтування цими принципами може призвести до втрати даних, порушення етичних норм обробки чутливої інформації та компрометації результатів. Тому хмарна інфраструктура потребує не лише технічного опанування, а й методологічного осмислення як частини наукової етики та практики.

Розглянуті безпекові патерни, зокрема принцип найменших привілеїв, ізоляція середовищ, централізоване логування, управління ролями та моделювання загроз, мають критичне значення не лише для забезпечення безпеки хмарної інфраструктури, а й для гарантування відтворюваності наукових досліджень. По-перше, ізоляція середовищ через окремі акаунти, підписки або проєкти дозволяє уникнути небажаного перехресного впливу між експериментами, забезпечує стабільність конфігурацій та чистоту результатів. Це особливо важливо для задач, які потребують високої точності, як-от: моделювання, машинне навчання, обробка чутливих даних. По-друге, управління доступом через тимчасові ролі (RVM, PIM, TTL) забезпечує контрольоване середовище виконання, де кожна дія має чітко визначені права та часові межі. Це сприяє трасуванню експериментальних дій, зменшує ризик втручання сторонніх компонентів і дозволяє точно відтворити умови дослідження. По-третє, логування та аудит (CloudTrail, Audit Logs, Access Analyzer) створюють повну історію змін і доступу до ресурсів, що є основою для верифікації результатів та аналізу впливу інфраструктурних факторів на перебіг експерименту. По-четверте, автоматизація через IaC та CI/CD дозволяє зберігати конфігурації в системах контролю версій, забезпечуючи повторюваність середовищ та можливість точного відтворення інфраструктури в будь-який момент часу.

Тим самим безпекові патерни, інтегровані в хмарну архітектуру, не лише знижують ризики, а й створюють технічно контрольоване, стабільне, прозоре середовище, яке відповідає вимогам наукової достовірності та відтворюваності.

Висновки. Проведене дослідження підтвердило, що хмарні платформи AWS, Azure та GCP забезпечують широкий спектр механізмів для реалізації безпекових патернів, необхідних у наукових дослідницьких середовищах. Водночас їхня ефективність залежить від правильного налаштування ролей, ізоляції середовищ та інтеграції з інструментами автоматизації. Порівняльний аналіз показав, що принцип найменших привілеїв, централізоване логування та моделювання загроз є критично важливими для запобігання витоку прав, прихованих точок доступу та перехресного впливу між середовищами. Відсутність цих механізмів створює ризики компрометації даних і порушення регуляторних норм. Використання IaC у поєднанні з CI/CD або GitOps-підходами забезпечує повторюваність, контрольованість і прозорість конфігурацій, що є

необхідною умовою для достовірності наукових експериментів. Практичні кейси підтвердили, що автоматизоване управління ролями (RVM в AWS, PIM в Azure, TTL-доступ у GCP) дозволяє зменшити ризики надмірного доступу та підвищити гнучкість адміністрування в мультихмарному середовищі.

Результати дослідження можуть бути використані для побудови безпечної, масштабованої та етично відповідної хмарної інфраструктури, що функціонує як повноцінний науковий інструмент, а не лише технічне середовище. Це відкриває перспективи для подальшої стандартизації хмарних архітектур у сфері комп'ютерних наук. Перспективи подальших досліджень лежать у площині розширення аналізу на інші хмарні платформи та гібридні моделі, а також у розробці освітніх модулів з DevSecOps для академічного середовища.

Список літератури

1. Dasari H. Infrastructure as Code (IaC) Best Practices for Multi-Cloud Deployments in Enterprises. *International journal of networks and security*. 2025. Vol. 05, Issue 01. P. 174-186. DOI: https://doi.org/10.55640/ijns-05-01-10.
2. Johnson O., Olamijuwon J., Cadet E., Osundare O., Samira Z. Designing multi-cloud architecture models for enterprise scalability and cost reduction. *Open Access Research Journal of Engineering and Technology*. 2024. Vol. 07(02). P. 101-113. DOI: https://doi.org/10.53022/oarjet.2024.7.2.0061
3. Thota R. Ch. Cloud-Native DevSecOps: Integrating Security Automation into CI/CD Pipelines. *International Journal of Innovative Research and Creative Technology*. 2024. Vol. 10. P. 1-19. DOI: https://doi.org/10.5281/zenodo.15036934
4. Трофименко О.Г., Дика А.І., Лобода Ю.Г. Аналіз інструментів тестування вебзастосунків. *Кібербезпека: освіта, наука, техніка.* 2023. № 4(20). С. 62-71. DOI: https://doi.org/10.28925/2663-4023.2023.20.6271.
5. Vidyasagar V. Multi-Cloud DevOps Automation for An Empirical Study on IaC, CI/CD, and Kubernetes Orchestration. *International Journal of Innovative Research in Education.* 2020. Vol. 01(01). P. 605-617. DOI: https://doi.org/10.5281/zenodo.15556322.
6. Bhat K., Prashanth K. Multicloud Orchestration using Terraform. *International Journal for Research in Applied Science and Engineering Technology.* 2022. DOI: https://doi.org/10.22214/ijraset.2022.44760.
7. Mulpuri G. Infrastructure as Code (IaC): Best Practices of Implementing IaC, Especially in Automating Infrastructure Provisioning and Management Using Terraform. *European Journal of Advances in Engineering and Technology*. 2023. Vol. 10(4). P. 56-62. URL: https://zenodo.org/records/11078219
8. Verdet A., Hamdaqa M., Da Silva L., Khomh F. Exploring Security Practices in Infrastructure as Code: An Empirical Study. *arXiv*. 2023. URL: https://arxiv.org/abs/2308.03952
9. Sonawane Y. Infrastructure as Code (IaC): Why It's a Game Changer in DevOps. URL: https://dev.to/yash_sonawane25/infrastructure-as-code-iac-why-its-a-game-changer-in-devops-1e1p
10. Michalowski M. Best Practices for Using Terraform: An In-depth Guide. *IEEE Computer Society*. 2024. URL: https://www.computer.org/publications/tech-news/trends/terraform-guide
11. Fuchs M. D. Policy as Code, Policy as Type. *arXiv*. 2025. DOI: https://doi.org/10.48550/arXiv.2506.01446
12. Трофименко О.Г., Пастернак Ю.Ю., Манаков С.Ю., Лобода Ю.Г. Автоматизація тестування вебсайтів електронної комерції. *Сучасна спеціальна техніка*. 2021. № 2(65). С. 46-59. DOI: https://doi.org/10.36486/mst2411–3816.2021.2(65).5.
13. Трофименко О.Г., Дика А.І., Лобода Ю.Г. Аналіз уразливостей та проблем безпеки вебзастосунків. *Системні технології.* 2023. № 3(146). С. 25-37. DOI: https://doi.org/10.34185/1562-9945-3-146-2023-03.

COMPARATIVE ANALYSIS OF SECURITY PATTERNS OF CLOUD PLATFORMS IN THE CONTEXT OF SCIENTIFIC REPRODUCABILITY

O. G. Trofymenko[1], P. O. Chykunov[1], D. Yu. Astakhov[1],
Yu. V. Molokanov, T. A. Farionova[2]

[1]National University "Odesa Law Academy"
23, Fontans'ka doroga st., Odesa, 65009, Ukraine
[2]Admiral Makarov National University of Shipbuilding
9, Geroiv Ukrainy Ave, Mykolaiv, 54007, Ukraine
Emails: trofymenko@onua.edu.ua, pavel@onua.edu.ua, astakhovnil@gmail.com,
molokanov9@gmail.com, tetyana.farionova@nuos.edu.ua

The study is devoted to a systematic analysis of the implementation of key security patterns in the cloud platforms Amazon Web Services, Microsoft Azure and Google Cloud Platform, considering their suitability for scientific research in the field of computer science. The relevance of the study is due to the need to ensure security, isolation, automation and reproducibility of the infrastructure in a multi-cloud environment. The purpose of the study is to systematize and comparatively analyze key security patterns in the cloud platforms AWS, Azure and GCP, considering their effectiveness, limitations and suitability for ensuring reproducibility, isolation and ethical compliance in scientific research environments. The research methodology is based on a systematic approach, which includes a comparison of cloud platform configurations based on practical cases, the use of static analysis tools (Checkov, tfsec) to assess the security of IaC code, modeling system behavior in failure scenarios and attacks to empirically verify the stability of the infrastructure. The article considers three practical cases: 1) for AWS – implementation of Role Vending Machine for automated assignment of temporary roles with limited rights, integration with GitHub Actions, Checkov and IAM Access Analyzer; 2) for Azure – multi-zone architecture with PIM, GitOps, ARM templates and network segmentation for fault tolerance modeling; 3) for GCP – use of Shared VPC, isolated projects, TPU and Vertex AI Experiments for NLP and deep learning tasks. Comparative analysis showed: the principle of least privilege significantly reduces the risks of rights leakage; the lack of logging and auditing creates "dead" rights that can be entry points for attacks; isolation of environments (through separate accounts, subscriptions, projects) is critical to prevent cross-influence; change management through CI/CD or GitOps provides repeatability and control; threat modeling allows you to identify vulnerabilities associated with resource placement; regulatory requirements (GDPR, encryption, geofencing) should be taken into account at the architectural planning stage. The research conducted allows us to conclude that a cloud infrastructure configured considering the principles of security, automation and ethical compliance can become a full-fledged scientific tool, and not just a technical environment. At the same time, neglecting these principles creates risks of data compromise, violation of regulatory norms and loss of reliability of results. The scientific novelty of the work lies in the comparative analysis of security models of three cloud platforms in the context of ensuring the reliability and ethical compliance of scientific experiments.

The practical significance of the results lies in the possibility of their use for building a secure, scalable and reproducible cloud infrastructure in scientific projects, in the professional activities of DevOps teams working with multi-cloud environments, as a basis for developing ethical and regulatory-compliant architectures in the field of computer science.

Keywords: cloud computing, multi-cloud environment, Infrastructure as Code, DevOps, DevSecOps, security patterns, CI/CD, GitOps, cloud architecture.

DOI 10.15276/imms.v15.no4.647
УДК 621.391

Informatics and Mathematical Methods in Simulation
Vol.15 (2025), No. 4, pp. 647-657

РОЗРОБКА ТА ДОСЛІДЖЕННЯ МЕТОДІВ ВИЗНАЧЕННЯ МАСИ ОБ'ЄКТІВ У РУСІ

Є. В. Шендрик[1], О. В. Головачова[2], В. В. Качеровський[3]

Національний університет «Одеська політехніка»
1, Шевченка пр., Одеса, 65044, Україна
Emails: e.v.shendryk@op.edu.ua[1], holovachova@op.edu.ua[2], 1562262@stud.op.edu.ua[3]

Розроблено метод підвищення точності вимірювання маси об’єкта при обмеженому часі зважування, який має високу точність та швидкодію за рахунок введення обмежень на величину отриманого значення постійної складової сигналу та використання формул заміни циклічних операцій. Запропоновані нові ефективні методи підвищення завадостійкості методу заданого діапазону частот шляхом його доповнення методами подвійного інтегрування вхідного сигналу або вагової функції, які знижують вплив високочастотних завад при автоматизованому вимірюванні маси рухомих об’єктів. Розроблено програмний модуль, який реалізує запропонований метод підвищення точності тензометричних вимірювань та може бути інтегрований в автоматизовані системи зважування об’єктів у русі при обмеженому часі спостереження.

Ключові слова: тензометричні вимірювання, тензодатчик, методи підвищення точності вимірювань, зважування у русі

Вступ. Динамічний розвиток автоматизованих систем управління та прогресивних інформаційних технологій у напрямку створення інтелектуальних засобів обробки інформації, зокрема, таких як методи побудови інформаційних моделей процесу, що автоматизується, обумовлює їх активне впровадження на всіх стадіях розробки й удосконалення автоматизованих систем вимірювання маси. Забезпечення автоматизації проведення процесу вимірювання, обробки отриманої інформації та подальше її накопичування, збереження й використання для задоволення інформаційних потреб залізничних або інших служб — основне призначення автоматизованих ваговимірювальних систем.

Автоматизований облік вантажопотоків, переміщуваних між регіонами країни або при перетинанні її державного кордону, є однією з найважливіших задач вирішуваних сьогодні. Відомо, що Україна володіє вельми розвиненою структурою залізничних магістралей, що зв’язують регіони країни, а також прикордонні держави, яких завдяки її геополітичному розташуванню виявляється досить багато. Тому своєчасне отримання вимірювальної інформації, що відбиває кількість перевезень, а також масу переміщуваного вантажу, формування єдиної інформаційної мережі вимірювань дозволить оптимально врегулювати напрямок перевезень, здійснити контроль втрат вантажів (особливо сипучих) при переміщуванні між станціями, а також, що є дуже важливим, задовольнити інформаційні потреби залізничних служб, які забезпечують контроль та цілісність залізничних магістралей.

Мета і задачі роботи. Скорочення часу зважування при незмінній точності вимірювань шляхом побудови інформаційної моделі процесу зважування. Для досягнення цієї мети передбачається вирішити такі задачі:

- проаналізувати методи і засоби вимірювання маси об’єктів у русі;
- розробити інформаційну модель оцінки маси об’єкту при обмеженому часі зважування;
- на базі інформаційної моделі реалізувати алгоритм вагового обліку.

Основна частина. Досліджено характеристики динамічних явищ, виникаючих у процесі зважування, на підставі чого одержана нова об'єктно-орієнтована модель сигналу отриманого з датчиків автоматизованої вагової платформи при проведені вагових вимірювань.

Доведено, що приведена об'єктно-орієнтована модель сигналу адекватна узагальненої моделі процесу зважування і представлена сукупністю трьох складових [1]

$$f(t) = D + A\sin(\Omega t + \psi) + \vec{\xi}(t), \quad (1)$$

де $f(t)$ — досліджуваний тензометричний сигнал;

D — постійна складова сигналу (інформативний параметр, відповідний масі об'єкта, що зважується);

$A\sin(\Omega t + \psi)$ — низькочастотна періодична складова сигналу;

$\vec{\xi}(t)$ — випадкова величина, що виникає під час зважування.

При цьому періодична завада представлена амплітудою — A, частотою — Ω і початковою фазою — ψ. Крім того, досліджуваний сигнал представлений сукупністю рівномірно розподілених у часі відрахунків $t_{i+1} - t_i = t_i - t_{i-1} = \Delta t$, де $i = \overline{0,n}$, $n+1 = N$ — кількість значень сигналу.

При довжині вагової платформи обмеженої 1,5 м, нижчої частотою коливань вагону 3 Гц та максимальній швидкості руху потягу 40 км/г тривалість вимірювального сигналу складає трохи більше чверті періоду низькочастотної періодичної складової сигналу. Тобто час спостереження сигналу — $T_н$ менше періоду завади $T_п$, $T_н < T_п$. Таким чином, розглядається випадок, коли періодична складова сигналу представлена менш чим одним періодом, а саме $T_н \approx \frac{1}{4} T_п$.

На підставі зазначених умов виміру побудована інформаційна модель, що відбиває не тільки характер поводження сигналу, але й оцінку внеску кожної із складових сигналу в результуючий вихідний сигнал.

Відповідно до цього задача була зведена до пошуку найбільш ефективного методу побудови інформаційної моделі процесу зважування, спрямованого на визначення оцінок інформативних параметрів приведеної об'єктно-орієнтованої моделі (1), зокрема, маси об'єкта — D, при заданих обмеженнях.

Дослідження показали, що особливої уваги заслуговують наступні методи:

– метод заданого діапазону частот;
– методи нелінійної регресії.

Показано, що *метод заданого діапазону частот* має високу точність оцінок досліджуваної моделі (1).

В основі цього методу лежить метод апроксимації узагальненим поліномом методом найменших квадратів [2-5]. Однак метод заданого діапазону частот різниться від вказаного тим, що має визначену систему базисних функцій, що відповідає моделі (1). Крім цього для відшукання інформативного параметру сигналу використовувався послідовний перебір частот із заданого діапазону. Метод базується на критерії найменших квадратів, тому найкращим вважається той результат, середньоквадратична похибка котрого прийме найменше значення.

Апроксимація узагальненим поліномом методом найменших квадратів, яка лежить в основі методу заданого діапазону частот, полягає у відшуканні серед поліномів m-го ступеня, $m \le n$

$$P_m(t) = a_0\phi_0(t) + a_1\phi_1(t) + \ldots + a_m\phi_m(t) \quad (2)$$

такого, для якого справедливий вираз

$$S = \sum_{i=0}^{n} (P_m(t_i) - f_i)^2 \to \min \tag{3}$$

де $a_0, a_1, \ldots, a_m$ — коефіцієнти узагальненого апроксимуючого полінома;
$\phi_0(t), \phi_1(t), \ldots, \phi_m(t)$ — задана система базисних функцій;
S — середньоквадратична похибка відхилення апроксимуючого полінома (2) від заданої функції $f(t)$.

Шукані коефіцієнти $a_0, a_1, \ldots, a_m$ полінома (2) визначаються із системи лінійних алгебраїчних рівнянь, вирішивши яку і підставивши знайдені значення $a_0, a_1, \ldots, a_m$ в (2), при відповідних базисних функціях, одержимо шуканий узагальнений апроксимуючий поліном.

У методі заданого діапазону частот проведення апроксимації здійснюється декілька разів, постійно змінюючи величину частоти заданих базисних функцій в межах заданого діапазону частот.

Застосувавши к моделі (1) тригонометричні формули $A \sin(\Omega t + \psi) = A \sin \psi \cos \Omega t + A \cos \psi \sin \Omega t$, де $A_1 = A \sin \psi = const$; $A_2 = A \cos \psi = const$, визначаємо систему базисних функцій $\phi_0(t), \phi_1(t), \ldots, \phi_m(t)$, обмеживши при цьому ступінь апроксимуючого полінома $m = 2$. При цьому зауважимо, що модель $D + A_1 cos(\Omega t) + A_2 sin(\Omega t)$ повністю адекватна моделі (1), різницею є тільки інший запис моделі (1), більш сприятливий для її дослідження запропонованим методом. Таким чином система базисних функцій приймає вид

$$\begin{cases} \phi_0(t) = 1; \\ \phi_1(t) = cos(\Omega t); \\ \phi_2(t) = sin(\Omega t). \end{cases} \tag{4}$$

Згідно (4), постійній складовій *D* відповідає функція $\phi_0(t)$, а періодична складова представлена сукупністю функцій $\phi_1(t), \phi_2(t)$.

Подальші обчислення зводилися до визначення коефіцієнтів a_0, a_1, a_2 таких, для яких величина (3) мінімальна. Алгоритм цих обчислень являє собою ітераційний процес і зводиться до побудови деякої множини $Z = \frac{\Omega_{max} - \Omega_{min}}{\Delta\Omega}$ апроксимуючих кривих у заданому частотному діапазоні $[\Omega max_{min}]$, із кроком $\Delta\Omega$.

На основі отриманих коефіцієнтів a_0, a_1, a_2 визначаються значення постійної складової сигналу і складових періодичної завади, припускаючи, що найкраща апроксимація ($S \approx 0$) була досягнута на *j*-й ітерації

$$D = a_{0_j}, A = \sqrt{a_{1_j}^2 + a_{2_j}^2}, \psi = arctg\left(\frac{a_{1_j}}{a_{2_j}}\right), \Omega = \Omega_{min}. \tag{5}$$

Проведено дослідження *методів нелінійної регресії*: метод лінеаризації моделі процесу (метод Гауса), метод Ньютона, метод Маркуардта і метод найшвидшого спуску, з метою виявлення ефективності їхнього використання до розв'язання поставленої задачі. Виявлено, що найкращим з розглянутих є метод Ньютона, який має найвищу точність, швидкість збіжності і найбільшу стійкість до вибору початкових умов у порівнянні з розглянутими методами. Показано, що ні метод Ньютона, ні інші методи нелінійної регресії не можуть застосовуватися для задачі проведення вагових вимірів при високій швидкості руху. Насамперед це зв'язано з тим, що розглянута модель сигналу має декілька локальних і тільки один глобальний мінімум функції середньоквадратичного відхилення (рис. 1), який приводить до точного результату. Як правило, в умовах обмеженого часу зважування, коли досліджуваний сигнал представлений чвертю періоду, пошук оптимальних оцінок приводить до влучення в область локального мінімуму, що в остаточному підсумку дає невірні результати і

вказує на неприйнятність використання методів нелінійної регресії для розв'язання поставленої задачі.

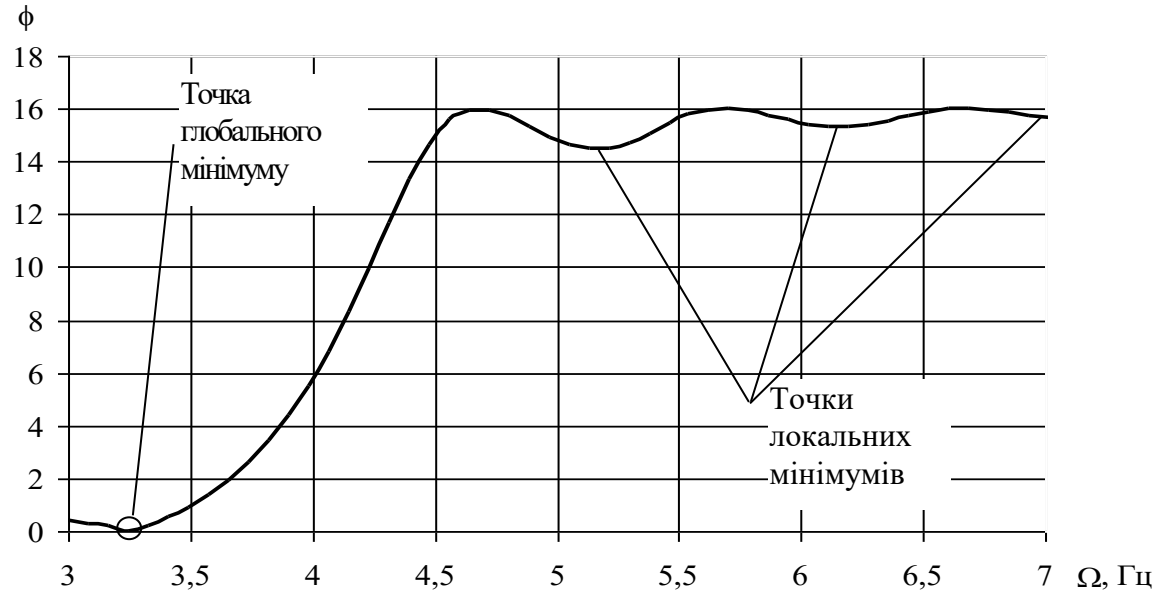


Рис. 1. Функція середньоквадратичного відхилення з одним глобальним і декілька локальних мінімумів

Приведено порівняльний аналіз методів (таблиця 1). Крім того, розглядається можливість використання кожного із методів для проведення вагових вимірів у різних умовах зважування і можливість їхнього впровадження у вже існуючі вагові комплекси з метою підвищення точності визначення інформативного параметру сигналу.

Таблиця 1.

Загальний порівняльний аналіз приведених методів

Найменування методу	Умови зважування	
	Похибка вимірювання при швидкості до 15 км/г, %	Похибка вимірювання при швидкості до 40 км/г, %
Звичайне усереднення	1,5	6
Метод заданого діапазону частот	0,1	0,7
Методи нелінійної регресії (метод Ньютона)	0,2	5

Відповідно до даних таблиці 1 зроблені висновки:

– методи нелінійної регресії мають великий потенціал при проведенні вагових вимірів із швидкістю руху об'єкта, що зважується, до 15 км/г, однак реалізація методів дуже складна, що вимагає їх доробки з ціллю забезпечення проведення вагових вимірів у реальному масштабі часу;

– метод заданого діапазону частот є єдиним прийнятним методом оцінки інформативного параметру сигналу при обмеженому часі зважування в реальних умовах, здатним визначати шукану величину, як при низькій, так і високій швидкості руху об'єкта, що зважується, забезпечуючи високу точність вимірювань.

Показано, що в деяких випадках похибка запропонованого методу може значно перевищити похибку, отриману шляхом усереднення значень сигналу. Встановлено, що причиною подібних випадків є несприятливий розподіл випадкового шуму, що у силу малої тривалості досліджуваного сигналу значно відхиляє апроксимуючу криву від реального значення представленого постійної складової і низькочастотною періодичною завадою, що у свою чергу приводить до помилкових оцінок параметрів.

Виявлено, що вплив випадкового шуму особливо гостро виявляється на граничних ділянках досліджуваного сигналу.

У якості методів підвищення завадостійкості методу заданого діапазону частот стосовно до поставленої задачі запропоновано використовувати:

– метод подвійного інтегрування;

– метод вагової функції;

– метод воріт.

Метод подвійного інтегрування [6] заснований на дослідженні реальних сигналів і припускає аналіз випадкового шуму $\vec{\xi}(t)$ у виді тригонометричного ряду

$$\vec{\xi}(t) = A_1 \sin(\Omega_1 t + \psi_1) + A_2 \sin(\Omega_2 t + \psi_2) + \ldots + A_m \sin(\Omega_m t + \psi_m) \quad (6)$$

у якому величини амплітуд A_1, A_2, ..., A_m як мінімум на порядок менше, а величини частот Ω_1, Ω_2, ... ,Ω_m як мінімум, на порядок більше відповідних величин сигналу (1).

Запропоновано двічі проінтегрувати сигнал (1), підставивши замість $\vec{\xi}(t)$ вираз (6). Інтегрування проводиться двічі по двох причинах. Однократне інтегрування недостатньо забезпечує подавлення випадкового шуму, і, як наслідок, не дає бажаного збільшення точності вимірів. При багаторазовому інтегруванні виникає похибка алгоритму чисельного інтегрування, що приводить до зниження точності одержуваного результату. Таким чином, двічі проінтегрований сигнал має вид

$$f(t)\,dtdt = \frac{Dt^2}{2} - \frac{A}{\Omega^2}\sin(\Omega t + \psi) - \frac{A_1}{\Omega_1^2}\sin(\Omega_1 t + \psi_1) - \ldots - \frac{A_m}{\Omega_m^2}\sin(\Omega_m t + \psi_m) \quad (7)$$

Як видно з (7), після дворазового інтегрування величина амплітуди кожної із складових сигналу ділиться на квадрат її частоти. З урахуванням того, що величини амплітуд шуму (6) на порядок менше, а величини частот на порядок більше відповідних величин сигналу (1), отримуємо істотне подавлення високочастотних складових сигналу. Подальший хід обчислень зводиться до використання методу заданого діапазону частот. При цьому пропонується нова система базисних функцій

$$\begin{cases} \phi_0(t) = t^2, \\ \phi_1(t) = cos(\Omega t), \\ \phi_2(t) = sin(\Omega t), \\ \phi_3(t) = t, \\ \phi_4(t) = 1. \end{cases}$$

Крім того, пропонується змінити спосіб обчислення оцінок шуканих параметрів сигналу (1):

$$D = 2a_{0_j},\ \Omega = \Omega_{min},\ A = \Omega^2\sqrt{a_{1_j}^2 + a_{2_j}^2},\ \psi = arctg\left(\frac{a_{1_j}}{a_{2_j}}\right).$$

Показано, що запропонований підхід дозволяє досягти десятикратного збільшення точності вимірів при малій тривалості досліджуваного сигналу.

Метод вагової функції [7] заснований на завданні вагових коефіцієнтів кожному із значень сигналу. Виявлено, що вплив випадкового шуму особливо гостро виявляється на граничних ділянках досліджуваного сигналу. Таким чином, для підвищення точності оцінок сигналу вводиться система вагових коефіцієнтів, відповідно до якої граничним значенням задається найменша вага, що збільшується в міру наближення до центральних значень досліджуваного сигналу.

Функцію, що визначає значення кожного вагового коефіцієнта w_i, визначена як

$$W(t) = \sin\left(\frac{\pi}{N}t + \frac{\pi}{2N}\right), \quad (8)$$

де N — кількість значень сигналу.

Нова система базисних функцій (4) приймає вид

$$\begin{cases} \phi_0(t) = W(t), \\ \phi_1(t) = cos(\Omega t)W(t), \\ \phi_2(t) = sin(\Omega t)W(t). \end{cases} \quad (9)$$

На величину (8) також збільшуються значення досліджуваного сигналу (1) — $f(t)W(t)$.

Показано, що існує ще один підхід, який дозволяє поліпшити точність оцінок параметрів тензометричного сигналу (1).

Шляхом множини експериментів та їхнього аналізу встановлено, що при застосуванні методу вагової функції обчислення середньоквадратичного відхилення в методі заданого діапазону частот повинне визначатися відповідно до формули

$$S = \sum_{i=0}^{N-1}(P_m(t_i) - f(t_i))^2 W(t_i). \quad (10)$$

Використання формули (10) дозволяє одержати мінімум середньоквадратичного відхилення відповідний найбільш точному значенню частоти із заданого діапазону $[\Omega_{min}\ldots\Omega_{max}]$, і, як наслідок, найбільш точне значення постійної складової сигналу D.

Показано, що запропонований підхід дозволяє досягти як мінімум десятикратного збільшення точності вимірів при малій тривалості досліджуваного сигналу, і, крім того, його використання дозволяє одержати більшу точність оцінок параметрів сигналу (1) у порівнянні з методом подвійного інтегрування.

Метод воріт. Встановлено, що в реальних умовах похибка вимірів при використанні кожного з розглянутих методів, у деяких випадках, може перевищити похибку, отриману шляхом простого усереднення сигналу, що є неприпустимим і може розцінюватися як збійна ситуація. У зв'язку з цим пропонується доповнити метод вагової функції, як кращий із представлених, таким чином, щоб гарантувати величину похибки не перевищуючу похибку усереднення значень сигналу при будь-яких умовах виміру.

Як таке доповнення запропоновано використовувати обмеження на вибір оцінок параметрів моделі (1) — ворота, які визначають припустимий діапазон отриманого значення постійної складової сигналу D. У цьому випадку використовуються дослідження реальних сигналів. З проведених досліджень відомо, що величина амплітуди низькочастотної складової сигналу (1) складає не більш 10...15 % величини постійної складової сигналу D. Таким чином, у результаті застосування методу заданого діапазону частот, обчислене значення величини $D = a_0$ не повинне перевищити середньоарифметичне значення досліджуваного сигналу $\overline{F}$, на величину

$$G = \pm\overline{F}/10, \quad (11)$$

де G — величина воріт.

Таким чином, з урахуванням уведених доповнень найбільш точним вважається той результат, що, по-перше, знаходиться в інтервалі

$$a_0 \in [\overline{F} \pm G], \quad (12)$$

де a_0 — поточний результат апроксимації, що відповідає D;

а по-друге, має найменшу середньоквадратичну похибку, розраховану по формулі (10). Якщо жодне з отриманих значень не задовольняє (12), то як шуканий результат виступає середньоарифметичне значення досліджуваного сигналу $\overline{F}$.

Проведено комп'ютерне моделювання з метою визначення точності оцінок параметрів сигналу, отриманих шляхом впровадження методу воріт (рисунок 2).

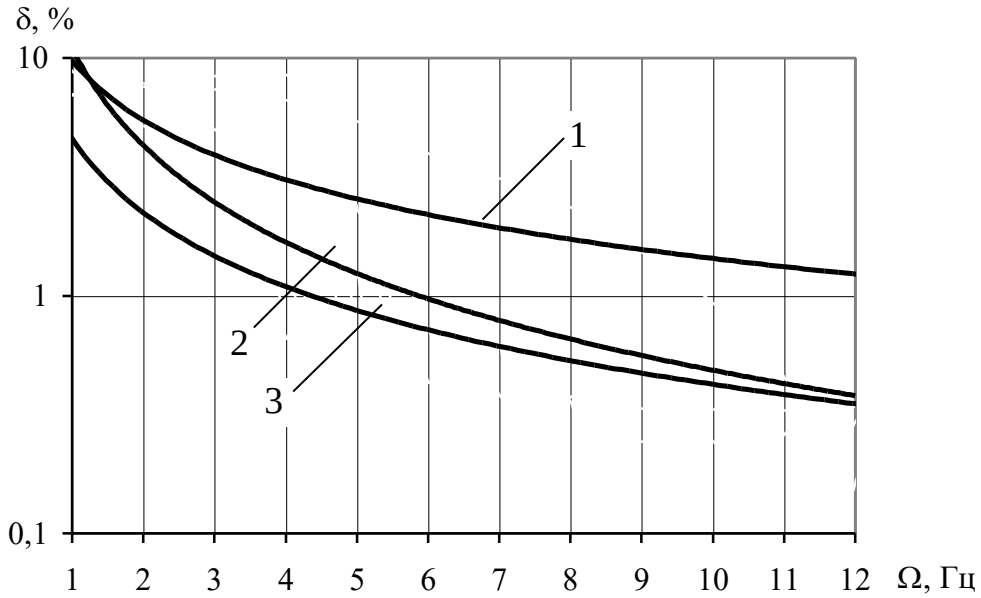


Рис. 2. Залежності відносної похибки виміру постійної складової сигналу від частоти, отримані: 1 – шляхом усереднення; 2 – з використанням методу вагової функції; 3 – з використанням методу вагової функції і методу воріт

Показано, що метод заданого діапазону частот удосконалений методом вагової функції і методом воріт (див. рисунок 2), є найкращим методом побудови інформаційної моделі оцінки маси об’єкта при обмеженому часі зважування, який дозволяє оцінити параметри об’єктно-орієнтованої моделі сигналу (1) при наявності збійних ситуацій у будь-яких умовах виміру.

Відповідно до проведених досліджень, реалізацію алгоритму побудови інформаційної моделі оцінки маси об’єкта при обмеженому часі зважування, необхідно робити на базі методу заданого діапазону частот із введенням у систему базисних функцій вагової функції і використанням методу воріт. При цьому необхідно забезпечити максимальну швидкодію обчислень шляхом застосування різних математичних перетворень, що зводять число обчислень до мінімально можливого значення.

У зв'язку з цим максимальним чином виключино з алгоритму циклічні обчислення, для чого використовуються отримані нами формули заміни циклічних обчислень.

В методі заданого діапазону частот, при формуванні системи лінійних рівнянь

$$\begin{aligned} c_{00}x_0 + c_{01}x_1 + c_{02}x_2 &= d_1, \\ c_{01}x_0 + c_{11}x_1 + c_{12}x_2 &= d_2, \\ c_{02}x_0 + c_{12}x_1 + c_{22}x_2 &= d_3, \end{aligned} \qquad (13)$$

де $c_{jk} = (\phi_j, \phi_k) = \sum_{i=0}^{n} \phi_j(t_i)\phi_k(t_i)$;

ліва частина рівняння, з урахуванням рівності діагональних коефіцієнтів та визначеної системи базисних функцій (9), обчислюється за допомогою формул заміни циклічних операцій.

Позначивши через q = sin(Ω), b = sin(ΩN), $p = sin\left(\frac{\pi}{N}\right)$, $r = sin\left(\frac{\pi}{2N}\right)$, e = cos(Ω), f = sin(ΩN), $g = cos\left(\frac{\pi}{N}\right)$, $h = cos\left(\frac{\pi}{2N}\right)$, де N — обсяг вибірки дискретних значень сигналу; обчислення коефіцієнтів рівняння (13) необхідно робити по наступним формулам

$$c_{00} = \frac{N}{2},\ c_{01} = \left(\frac{(2ghrp(f-1)-qbh^2)(1-e)+qb(g^2-e)}{2(2g^2(1-e)+e^2-1)} + \frac{h^2(f-1)+1-f}{2}\right),$$

$$c_{02} = \left(\frac{q(f-1)(e-g^2+h^2(1-e))+2gbhrp(1-e)}{2(2g^2(1-e)+e^2-1)} + \frac{b(h^2-1)}{2}\right),$$

$$c_{11} = \left(\frac{N}{4} + \frac{f^2(h^2-1)-h^2+1}{2} + \frac{rphg(f^2-1)}{2(g^2-e^2)} + \frac{fqbe(2e^2-g^2-1-2h^2(e^2-1))}{4(g^2-e^2)(e^2-1)}\right),$$

$$c_{12} = \left(\frac{fb(h^2-1)}{2} + \frac{qe(f^2-1)(1-2e^2+g^2)}{4(g^2-e^2)(e^2-1)} + \frac{h^2qe(f^2-1)+bhfrpg}{2(g^2-e^2)}\right),$$

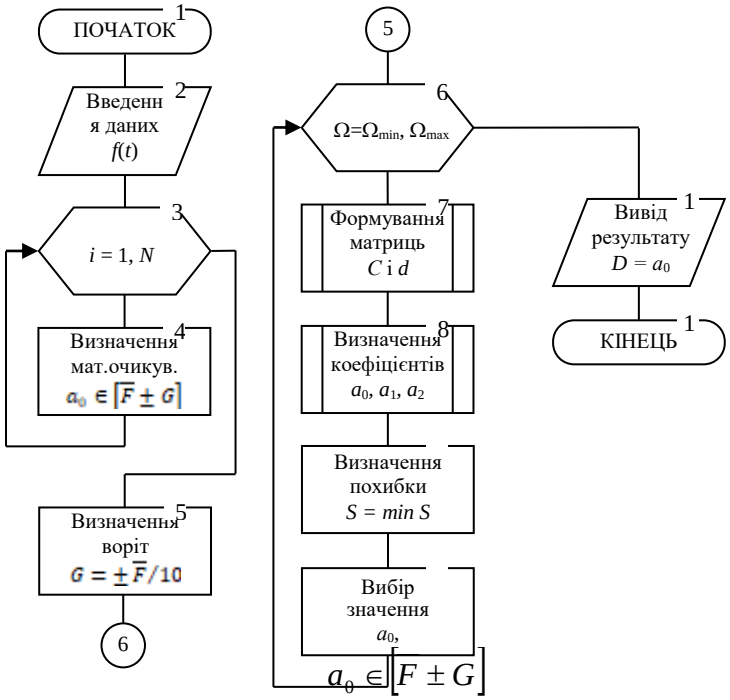
$$c_{22} = \left(\frac{N}{4} - \frac{f^2(h^2-1)-h^2+1}{2} - \frac{rphg(f^2-1)}{2(g^2-e^2)} - \frac{fqbe(2e^2-g^2-1-2h^2(e^2-1))}{4(g^2-e^2)(e^2-1)}\right).$$

При цьому обчислення правої частини рівняння (13) залишається незмінним — $d_k = \sum_{i=0}^{n} f(t_i)\phi_k(t_i)\,W^2(t_i)$, $k = 0, 1, 2$.

Слід зазначити, що використання зазначених формул дозволяє уникати виконання циклічних операцій для розрахунку коефіцієнтів лівої частини рівняння. При цьому по кількості виконуваних операцій зазначені формули вимагають меншої кількості часу обчислень, тому що із збільшенням кількості значень досліджуваного сигналу час обчислень коефіцієнтів лівої частини рівняння (13) залишається незмінним.

Крім того, слід зазначити, що існують дві виняткові ситуації, коли застосування зазначених формул виявляється неможливим, тому що в цьому випадку в деяких формулах знаменник виявляється рівним нулю. При цьому для ситуацій $\Omega = \frac{\pi}{N}$ і $\Omega = \frac{2\pi}{N}$ відповідно використовуються інші формули заміни циклічних операцій.

Таким чином, у результаті проведених експериментів було встановлено, що час обчислення проміжного результату за допомогою запропонованих формул зменшується, приблизно, в 10 разів в порівнянні із звичайним чином обчислення, що



цілком задовольняє проведенню вимірів у реальному масштабі часу. Алгоритм запропонованого методу представлений на рисунку 3.

Рис. 3. Алгоритм побудови інформаційної моделі оцінки маси об'єкта при обмеженому часі зважування, застосовуваний у системах вагового обліку

Проведено натурні експерименти, що підтверджують доцільність використання запропонованого рішення для задачі оцінки маси об’єкта при обмеженому часі зважування (рисунок 4). При цьому для приведених графіків (див. рисунок 4) швидкість руху потягу складала 40 км/г.

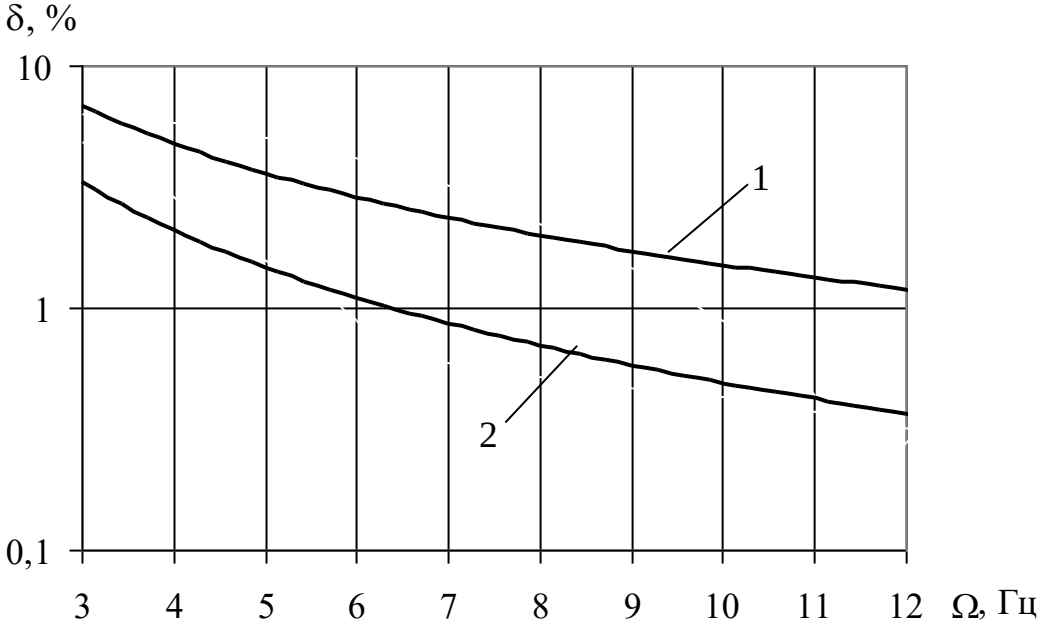


Рис. 4. Залежності відносної похибки виміру постійної складової сигналу від частоти, отримані: 1 — шляхом усереднення; 2 — з використанням запропонованого алгоритму

Зазначено, що програмна реалізація приведеного алгоритму в складі автоматизованого ваговимірювального комплексу при обмеженому часі зважування дозволяє підвищити точність вимірів у 2...7 разів стосовно усереднення значень сигналу, що на сьогоднішній день є кращим показником точності в рамках поставленої задачі.

Висновки. Основні наукові і практичні результати роботи полягають у наступному:

– проаналізовано сучасні методи та засоби для задачі підвищення точності та швидкодії автоматизованих ваговимірювальних систем. Доведено, що класичними методами та засобами, заснованими на усередненні вибірки сигналу не можливо забезпечити необхідну точність вимірювань, тому що їхнє використання не враховує особливості вхідного сигналу при автоматизованому зважуванні об’єктів на високій швидкості руху. Цей факт підтверджує доцільність розробки нових методів обробки сигналу, таких як методи побудови інформаційної моделі оцінки маси об’єкта при обмеженому часі зважування.

– досліджено автоматизовану ваговимірювальну систему з метою розробки нової моделі автоматизованого процесу зважування об’єктів у русі як базової моделі для подальших розробок методів побудови інформаційної моделі для оцінки маси об’єктів, що рухаються з підвищеною швидкістю.

– на базі нової моделі зважування вагонів у русі розроблено метод заданого діапазону частот, який шляхом послідовного перебору частот із заданого діапазону визначає оцінки інформативних параметрів тензометричних сигналів для побудови інформаційної моделі процесу автоматизованого зважування об’єктів у русі.

– запропоновані нові ефективні методи підвищення завадостійкості методу заданого діапазону частот шляхом його доповнення методами подвійного інтегрування вхідного сигналу або вагової функції, які знижують вплив високочастотних завад при автоматизованому вимірюванні маси рухомих об’єктів, поліпшуючи інформаційні параметри створюваної інформаційної моделі.

– з урахуванням методу вагової функції розроблено метод та алгоритм побудови інформаційної моделі оцінки маси об’єкта при обмеженому часі зважування, який має високу точність та швидкодію за рахунок введення обмежень на величину отриманого значення постійної складової сигналу — метод воріт, та використання формул заміни циклічних операцій.

Список літератури

1. Копытчук Н.Б., Огинский В.Н., Милейко И.Г. Оценка информативных параметров сигналов на фоне помех при ограниченном времени наблюдения. *Тр. Одес. политехн. ун-та.* 1999. Вып. 3(9). С. 149 — 152.
2. Фиакко А., Мак-Кормик Г. Нелинейное программирование. Методы последовательной безусловной минимизации. М.: Мир, 1972. 240 с.
3. Демидович Б.П., Марон И.А., Шувалова Э.З. Численные методы анализа. Приближение функций, дифференциальные и интегральные уравнения. М.: Наука, 1967. 368 с.
4. Копытчук Н.Б., Шендрик Е.В. Использование метода наименьших квадратов для оценки параметров сигнала с периодической помехой при ограниченном времени наблюдения. *Тр. Одес. политехн. ун-та.* 1999. Вып. 3(9). С. 167 — 169.
5. Демидович Б.П., Марон И.А. Основы вычислительной математики. М.: Наука, 1970. 664 с.
6. Копытчук Н.Б., Шендрик Е.В. Повышение точности метода наименьших квадратов посредством интегрирования. *Праці міжнародної конференції з управління “Автоматика”*. *Одеса*. 2001. С. 77 — 78.
7. Шендрик Е.В. Повышение эффективности алгоритма метода наименьших квадратов путем введения модифицированного метода весовой функции. *Праці УНДІРТ*. 2002. № 1 (29). С. 88 — 90.

DEVELOPMENT AND RESEARCH OF METHODS FOR DETERMINING THE MASS OF OBJECTS IN MOTION

E.V. Shendryk[1], O.V. Golovachova[2], V.V. Kacherovskiy[3]

National Odessa Polytechnic University
1, Shevchenko Ave., Odesa, 65044, Ukraine
Emails: e.v.shendryk@op.edu.ua[1], holovachova@op.edu.ua[2], 1562262@stud.op.edu.ua[3]

The work develops a method for improving the accuracy of object mass measurement with limited weighing time, which has high accuracy and speed due to the introduction of restrictions on the value of the obtained constant component of the signal and the use of formulas for replacing cyclic operations. New effective methods are proposed to improve the noise immunity of the specified frequency range method by supplementing it with methods of double integration of the input signal or weight function, which reduce the influence of high-frequency noise during automated measurement of the mass of moving objects. A software module has been developed that implements the proposed method for improving the accuracy of strain gauge measurements and can be integrated into automated systems for weighing objects in motion with limited observation time.
Keywords: strain gauge measurements, strain gauge, methods for improving measurement accuracy, weighing in motion

ІНФОРМАТИКА ТА МАТЕМАТИЧНІ МЕТОДИ В МОДЕЛЮВАННІ

Том 15, номер 4, 2025. Одеса – 194 с., іл.

INFORMATICS AND MATHEMATICAL METHODS IN SIMULATION

Volume 15, No. 4, 2025. Odesa – 194 p.

Засновник: Національний університет «Одеська політехніка»

Зареєстровано Міністерством юстиції України 04.04.2011р.
Свідоцтво: серія КВ № 17610 - 6460Р

Друкується за рішенням Вченої ради Національного університету «Одеська політехніка», (протокол №7 від 09.12.2025р.)

Адреса редакції: Національний університет «Одеська політехніка»,

1, Шевченка проспект, Одеса 65044 Україна

Web: www.immm.op.edu.ua (immm.opu.ua)

Email: immm.ukraine@gmail.com