ISSN 2224-087X

ЕЛЕКТРОНІКА ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

ELECTRONICS
AND INFORMATION TECHNOLOGIES

Збірник наукових праць

Випуск 30

2025

ELECTRONICS AND INFORMATION TECHNOLOGIES

ЕЛЕКТРОНІКА ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

Issue 30

Scientific journal

Published 4 issue per year

Published since 1966

Випуск 30

Збірник наукових праць

Виходить 4 рази на рік

Видається з 1966 р.

Ivan Franko National University of Lviv

Львівський національний університет імені Івана Франка

2025

ЗАСНОВНИК: ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ІВАНА ФРАНКА

Друкується за ухвалою Вченої Ради
Львівського національного університету
імені Івана Франка
протокол №85/06 від 25.06.2025 р.

Реєстрація суб'єкта у сфері друкованих медіа:
Рішення Національної ради України з питань
телебачення і радіомовлення № 1877 від
30.05.2024 р. Ідентифікатор медіа R30-04912

У 1966–2010 рр. збірник виходив під назвою «**Теоретична електротехніка»**

Збірник «Електроніка та інформаційні технології» містить оригінальні результати досліджень з електронного матеріалознавства, моделювання фізичних явищ, процесів і систем електроніки, обробки сигналів і зображень, інформаційних технологій.

“Electronics and information technologies” journal contains original research results on electronics material science, modelling of physical phenomena, processes and electronic systems, signal and image processing and information technologies.

Адреса редакційної колегії:
Львівський національний університет імені Івана Франка, факультет електроніки та інформаційних технологій, вул. Ген. М. Тарнавського, 107, 79017, Львів, Україна
тел. (+38) (093) 864-01-19

Editorial office address:
Ivan Franko National University of L'viv,
Faculty of Electronics and Computer Technologies
107 Tarnavsky St., UA–79017,
Lviv, Ukraine
tel. (+38) (093) 864-01-19

e-mail: elit@lnu.edu.ua
web-сайт: http://publications.lnu.edu.ua/collections/index.php/electronics/index

АДРЕСА РЕДАКЦІЇ, ВИДАВЦЯ І ВИГОТОВЛЮВАЧА:

Львівський національний університет імені Івана Франка
вул. Університетська, 1, 79000 Львів, Україна

Свідоцтво про внесення суб'єкта видавничої справи до Державного реєстру видавців, виготівників і розповсюджувачів видавничої продукції. Серія ДК № 3059 від 13.12.2007 р.

CONTENTS

INFORMATION SYSTEMS AND TECHNOLOGIES

MODELING OF PROCESSES AND EFFECTS

MATERIALS FOR ELECTRONIC ENGINEERING

ЗМІСТ

ІНФОРМАЦІЙНІ СИСТЕМИ ТА ТЕХНОЛОГІЇ

МОДЕЛЮВАННЯ ПРОЦЕСІВ ТА ЯВИЩ

МАТЕРІАЛИ ЕЛЕКТРОННОЇ ТЕХНІКИ

ISSN 2224-088X (print) • ISSN 2224-0888 (on-line)

Electronics and Information Technologies, 2025, 30, 5–14
http://publications.lnu.edu.ua/collections/index.php/electronics/index

UDC 004.89

DESIGN AND IMPLEMENTATION OF AN INTELLIGENT SEARCH SYSTEM BASED ON NEURAL NETWORKS

Vitalii Bozhovskiy **, *Halyna Klym*** ***
Lviv Polytechnic National University,
12 Stepan Bandera St., Lviv 79013, Ukraine

Bozhovskiy, V., Klym, H. (2025). Design and implementation of an intelligent search system based on neural networks. *Electronics and Information Technologies*, 30, 5–14.
https://doi.org/10.30970/eli.30.1

ABSTRACT

Background. In the digital information era, the ability to retrieve relevant data quickly and accurately is increasingly critical. Traditional search engines such as Google or Bing rely on keyword matching, which can fail in cases of vague queries, multilingual content, or media-based searches. The rapid development of neural networks and AI technologies introduces new opportunities to enhance search systems by understanding context, semantics, and user behaviour. This study aims to develop a search system based on ElasticSearch, integrating multiple neural network modules to improve search precision, personalisation, and flexibility.

Methods. The proposed system includes four main components: ElasticSearch for full-text indexing, a convolutional neural network for image recognition, a graph-based semantic model for query expansion, and a ranking model based on historical user interactions. The backend is developed in Python using Visual Studio, with modular AI components that can be activated or disabled by the user. The semantic model represents terms as graph nodes and semantic proximity as weighted edges, enabling dynamic context-driven query refinement. Additional features include synonym detection, citation filtering, and user-specific ranking.

Results and Discussion. Two key experiments were performed. The first examined system performance by testing search speed across database sizes ranging from 100 to 100,000 records. It was found that even with all neural modules enabled, latency remained minimal, confirming system scalability. The second experiment assessed the impact of training data on the quality of the semantic model. A model trained on low-quality, AI-generated data resulted in incoherent word associations and poor query expansion. In contrast, a model built on human-curated texts produced clear, logical semantic links and significantly improved search relevance. The image search function demonstrated the system's ability to identify relevant visual content based on vague or partial user input, while the context expansion model enhanced result diversity and accuracy even with incomplete or ambiguous queries.

Conclusion. This work presents a hybrid search engine that effectively integrates traditional indexing with AI-powered features. The system offers robust text and image search capabilities, intelligent semantic understanding, and personalised ranking. Experiments confirmed its efficiency, relevance, and adaptability across varying data conditions and resource levels. With modular architecture and advanced context handling, the system addresses limitations of conventional search engines and sets a strong foundation for future development in intelligent information retrieval.

Keywords: Search system, search accuracy, neural networks, ElasticSearch.

INTRODUCTION

Searching for information has become an essential part of modern life. Every day, users unknowingly interact with search engines, whether by searching for documentation on Google when writing code, browsing photos through social media tags, or retrieving emails by keywords [1,2]. With the rapid advancement of artificial intelligence (AI), systems like ChatGPT and Bard [3,4] have emerged, further changing the way we search and interact with information. These AI-based platforms do not simply return search results but engage in semantic understanding, enhancing user experience.

Search engines, while designed with the same overarching goal of retrieving relevant information, differ significantly in their functionality and design. For example, Google helps users find websites, Google Maps helps users navigate to physical locations, and social media search engines focus on locating individuals or tagged content [5]. These differences dictate distinct algorithmic approaches and design choices for each system. Unlike traditional search engines, modern AI systems treat queries as questions posed to a human, offering responses that reflect deeper semantic understanding [6].

Traditional search engines often rely on exact word matches, but this approach is highly impractical in real-world scenarios. For example, users may make spelling errors, use synonyms, or express intent imprecisely, leading to poor results when exact matching is applied. To address this, full-text search engines, such as ElasticSearch, incorporate context, enabling systems to account for minor errors or alternative word choices and improve accuracy. This method improves the search process but still faces challenges when dealing with ambiguous or incomplete queries [7]. Furthermore, global search engines must account for variations in user backgrounds, such as language or cultural expectations, which may influence the results even for the same query [8].

Another pressing challenge is scalability, as information continues to grow exponentially. Search engines must be able to process massive volumes of data without sacrificing speed or accuracy. For this, advanced neural network models, such as convolutional neural networks (CNNs) for image recognition or recurrent neural networks (RNNs) for semantic understanding, are increasingly employed. These models must work efficiently with large datasets, enabling real-time processing of user queries, all while maintaining a high level of personalisation and relevance [9,10].

Additionally, users increasingly expect personalised results, meaning that search systems must not only retrieve relevant content but also learn and adapt to individual preferences over time. This involves integrating data from user interactions and feedback to refine ranking algorithms and query results. While such personalisation has greatly improved search performance, it also raises privacy concerns, which need to be addressed through proper data protection and ethical design [11,12].

Despite these advances, the development of effective search engines remains complex, as these systems must contend with linguistic diversity, multimedia content, varying user expectations, and large-scale data processing. As the demands on search engines continue to evolve, there is a pressing need for systems that can intelligently process and retrieve information while adapting to the needs of individual users [13,14].

The goal of this study is to conduct an in-depth analysis of existing search systems, identifying the tools, techniques, and neural technologies they use. Through this analysis, we aim to identify key challenges and solutions, as well as design and implement a custom search engine tailored for a social network project. Our system will incorporate full-text search capabilities, neural modules for semantic expansion, image-based retrieval, and personalised ranking. The system's performance, scalability, and user experience will then be evaluated through experimental testing to identify improvements over existing methods [15].

METHODS

Designing a search system. In this study, we present a search system that goes beyond the functionality of a traditional search engine by incorporating advanced neural network technologies. The system is designed to provide a more dynamic and user-centric experience by integrating several key components: the search engine itself, a multi-level graph neural network, an image analysis neural network, and a neural network responsible for ranking results based on user behaviour. This section describes the design of the system, emphasizing the interaction between these components and their role in optimising the search process.

Search Engine component is responsible for retrieving results from the database based on user queries. A multi-level graph neural network that processes user queries, refines and expands them by considering the semantic relationships between terms in the query. An image analysis neural network is designed to analyse and interpret images, extracting relevant information that is subsequently stored in the database for future searches. Ranking neural network trained on historical user data that ranks search results based on past interactions and user preferences. These components work in tandem to enhance the accuracy and relevance of search results. The flow of information between these subsystems, as well as the user's role in the interaction, is illustrated in Fig. 1.

The user interacts with the system in a series of steps that involve multiple stages of data processing and refinement. Initially, the user creates a post that may contain text, images, or both. If an image is included, the image analysis neural network is triggered to analyse the content of the image, extracting relevant features and storing them in the system's database. It is important to note that at this stage, the image analysis neural network operates independently of the search engine, performing its task of data extraction.

Once the data is stored, the system is ready to process subsequent user queries. If the query is image-based, it is directly sent to the search engine, which searches the database and returns relevant image results to the user. In the case of a text-based query, the process is more complex. The query is first passed to the multi-level graph neural network, which analyses and refines the query to improve its semantic accuracy. Rather

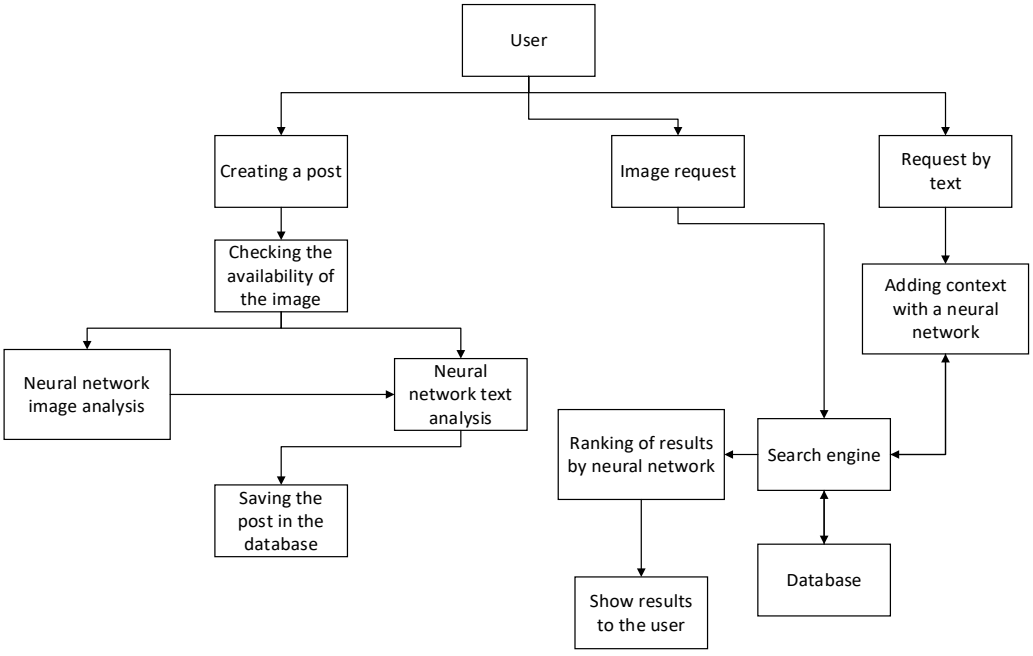


Fig. 1. The process of user interaction with the system.

than simply transmitting the modified query, this neural network actively interacts with the query to ensure it is more precise, helping the search engine to retrieve more relevant results.

Once the search engine returns the results, the ranking neural network takes over. It evaluates the results based on historical data, ranking them according to the user's past behaviour and preferences. The ranked results are then presented to the user. This approach allows the system to provide results that are not only accurate but also personalised, based on the individual user's search history and interests.

A key advantage of our system is its flexibility, allowing users to control the extent to which neural networks are involved in the search process. At any point in the system where neural networks are used, users have the option to disable them. This feature ensures that users can choose a more traditional, non-AI-driven search experience if they prefer. In some cases, users may find that the neural networks' refinement of queries and ranking of results do not align with their expectations, and the ability to turn these features off provides a tailored experience. This level of customisation is a significant advantage over other search systems, which often offer limited control over the search process.

Fusion of information for a multisensory system. The next phase of development for our search system involves the implementation of a neural network model designed to enhance the user's search query. The primary task of this model is to expand the query by utilising a range of techniques aimed at improving the accuracy and breadth of search results. The expansion process follows a systematic approach. The neural network begins by analysing the user's original query. If the initial query results in very few matches or limited synonyms, the model attempts to generate and integrate synonyms to improve the query's reach. Based on the newly generated words or phrases, the model expands the query and sends it to the search engine for further refinement and result retrieval. If the model identifies corresponding words or synonyms within the system, it ensures that relevant results are returned to the user.

This process works seamlessly when the user does not manually alter the query. For example, in a query like "ignorant [girl]," the neural network identifies the word "girl" within brackets and uses it as a cue for modification. The system then searches for syntactically close words to refine the query further. This enables the model to send new queries to the search engine until it produces the most accurate and relevant results, if possible.

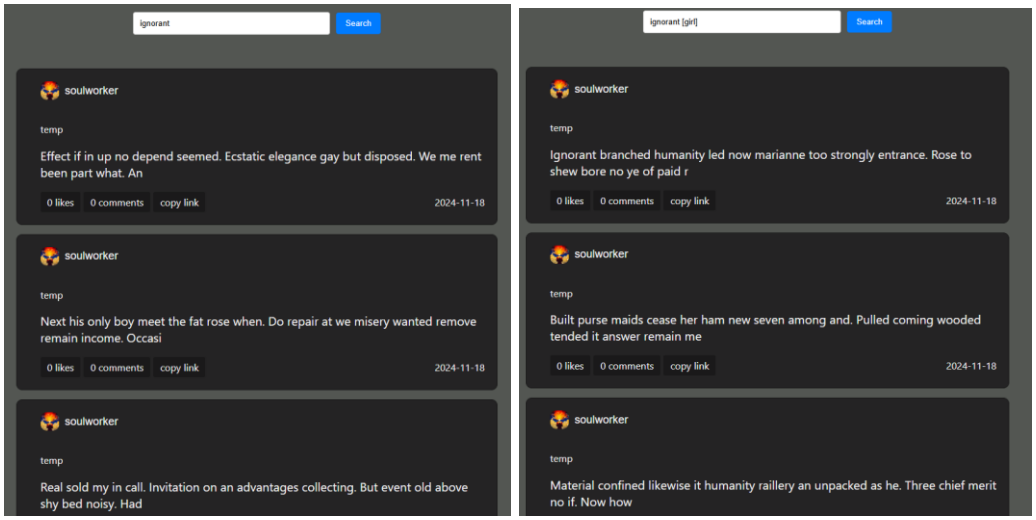
The core of this neural network is based on representing words as vectors within a multi-dimensional space, where semantically similar words are closer to each other. This is achieved by considering the words as nodes within a graph, where edges connect semantically related words, and each edge is assigned a numerical value that quantifies the proximity of those words in the semantic space. By leveraging these connections, the model can generate expanded search queries with a higher likelihood of returning accurate and relevant results. Fig. 2 illustrates how the neural network operates.

These modifications enhance the model’s ability to find more relevant results by broadening the scope of the original search while maintaining the semantic integrity of the query.

To provide users with more control over the search process, we offer the option to disable the neural network model. By using the “-nomodel” modifier, users can opt out of the query expansion process and return to a more traditional search experience. This feature is valuable for users who may prefer a simpler search approach or believe that the neural network's modifications might not always align with their specific needs.

RESULTS AND DISCUSSION

Impact of Data Quality on Model Performance. During the development of the search system, a custom neural network was created and trained to enhance query interpretation and improve search precision. One of the most critical factors influencing the effectiveness of such a model is the quality of the training data. The performance of neural



(a) (b)

Fig. 2. An example of a search query without modification(a), where the search engine processes the query as it is presented by the user and the query after modification (b), where the neural network has expanded and refined the original request by adding semantically related terms and synonyms.

networks, particularly in natural language processing tasks, is heavily dependent not only on the quantity but more importantly on the relevance and coherence of the data. Leading organisations in the field, such as OpenAI, emphasise the importance of high-quality user-generated content for training large-scale language models.

To empirically assess the role of data quality, an experiment was conducted involving the training of two distinct models using datasets of varying quality. The first model was trained on a large dataset consisting of randomly selected, low-quality texts, many of which were generated by other neural networks. Although the dataset comprised several thousand samples, the training process yielded suboptimal results. The semantic graph produced by this model, presented in Fig. 3a, reveals a high degree of disorder. Most nodes (words) are connected with high-weight edges, indicating an artificial and misleading semantic proximity between unrelated terms. This lack of structural coherence renders the model ineffective for search-related tasks, as it fails to capture meaningful linguistic relationships.

In contrast, the second model was trained on a smaller but carefully curated corpus of texts authored by real users. Despite the inherent variability in user-generated content, the selection process prioritised syntactic clarity, topical relevance, and contextual consistency. The resulting semantic graph (Fig. 3b) displays a logical and interpretable network of word associations. Connections between terms are no longer arbitrary, reflecting a semantically sound structure suitable for query expansion and refinement in the search process.

This experiment demonstrates that data quality significantly outweighs quantity when training models for semantic tasks. Excessive data without proper filtering may introduce noise, reduce model robustness, and impair downstream performance. Furthermore, these findings help explain why generative neural networks often struggle to improve through self-generated content, which lacks the depth and structure of well-formed human language. Accordingly, the developed search system integrates the second model, as it provides a stable semantic foundation aligned with real user needs and expectations.

System Quality and Performance Testing. To evaluate the efficiency and scalability of the developed search system, a performance testing experiment was conducted under controlled conditions. Before testing, a utility was implemented to populate the database with synthetic entries, and computational resources were intentionally constrained to

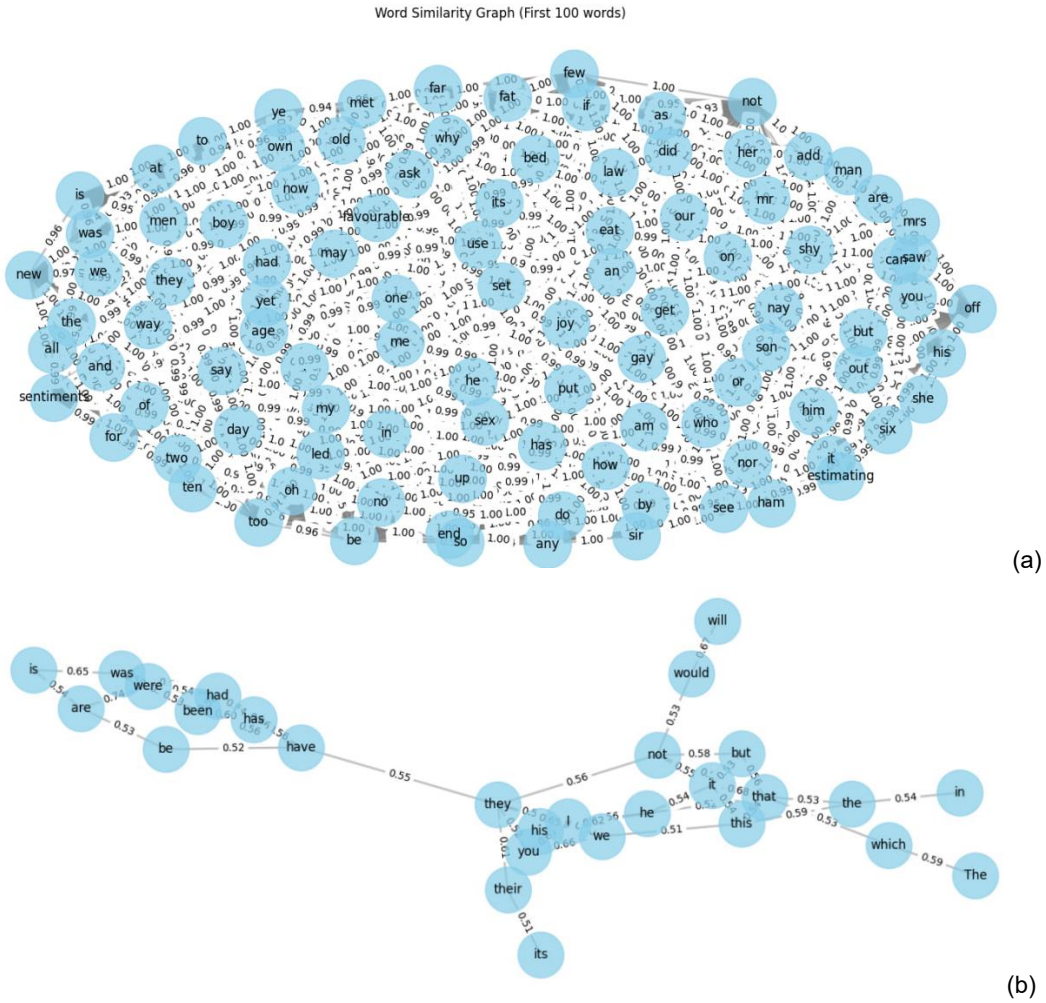


Fig. 3. Graph structure derived from low-quality (a) and high-quality (b) training data.

simulate realistic usage conditions. The objective was to assess the system's response time when processing search queries across varying dataset sizes.

Four distinct dataset sizes were selected for the experiment: 100, 1,000, 10,000, and 100,000 records. These quantities were deemed sufficient to demonstrate performance trends, especially under resource-limited conditions. Three types of search operations were tested: a standard text query, a query involving image analysis and a query processed through all integrated neural networks. The execution times for each scenario were recorded and visualised in the resulting performance graph (Fig. 4).

As illustrated in Fig. 4, standard queries consistently demonstrated the fastest response times. The most noticeable change in latency occurs between the dataset sizes of 1,000 and 10,000 records. Beyond this point, performance degradation becomes more gradual. Although queries involving image analysis and neural network processing introduce some additional delays, the increase in execution time remains within acceptable margins, amounting to only a few milliseconds per additional increment of records.

In addition to efficiency, the quality of search results was also tested and analysed. Used a benchmark set of 50 representative queries. These queries were executed in all three search modes, and standard information retrieval metrics were computed to assess

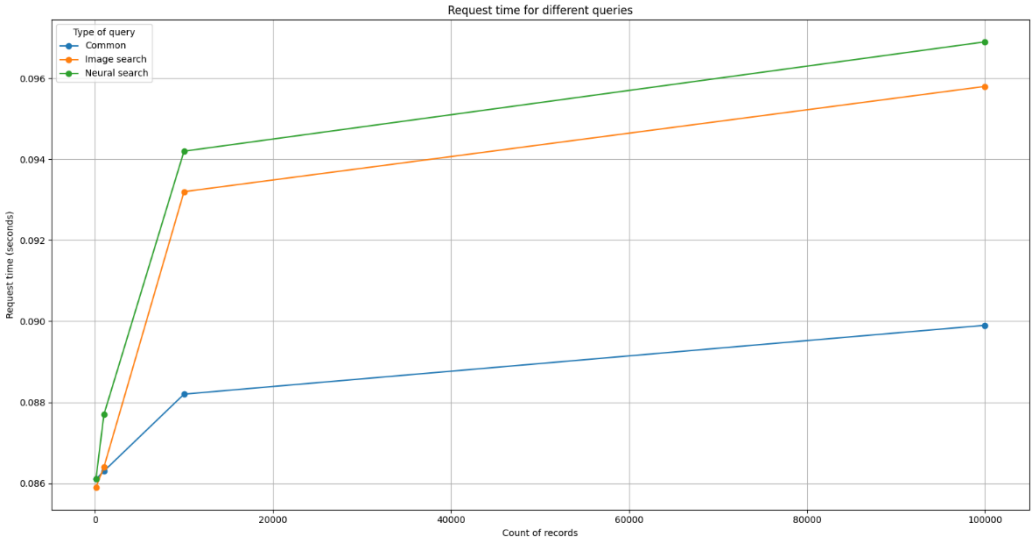


Fig. 4. Requested time as a function of the number of records.

rankings quality: Mean Reciprocal Rank (MRR) and Normalised Discounted Cumulative Gain at rank 10 (NDCG@10). The results are presented in Table 1.

As represented in Table 1, the neural search significantly increased quality performance compared to common search across metrics, which indicates improving relevance and ranking of search results. As a result, this test shows that neural search improves the user experience with the search engine.

These results suggest that while the inclusion of neural networks marginally affects search latency, it does not significantly hinder user experience. The search engine remains responsive and efficient, even as the volume of data increases. This outcome indicates that the system architecture and neural integration are sufficiently optimised to handle large-scale queries without compromising performance.

Table 1. **Quality metrics representation.**

Metric	Common search	Image Search	Neural Search
MRR	0.55	0.6	0.76
NDCG@10	0.53	0.58	0.74

CONCLUSION

As a result of the research conducted, an intelligent search system was developed based on modern technologies for information processing and analysis, in particular, artificial neural networks. At the initial stage, a comprehensive analysis of the search engine market was carried out, which revealed the main shortcomings of existing solutions. Special attention was paid to persistent issues that remain unresolved by current systems, which substantiated the need to create a new tool capable of addressing these challenges.

During the research, the key aspects that should be considered when developing a custom solution were identified. A suitable technological framework was selected, enabling the implementation of a search system with extended functionality. A central feature of the system is the integration of a proprietary neural network that refines user queries and improves result relevance. Additionally, the system provides the ability to activate or deactivate individual modules, ensuring adaptability to diverse user needs.

Upon completion of the development phase, a series of experiments was conducted to evaluate the performance and effectiveness of the system. Testing included measuring technical performance, analysing the impact of data quality and volume on model structure, and assessing the contribution of neural networks to query refinement. The results showed that high-quality and semantically rich data lead to logical and useful relationships within the model, while large volumes of low-quality data result in chaotic and less effective knowledge structures. It was also demonstrated that neural networks significantly enhance search relevance while only minimally affecting request processing speed, indicating a high degree of system optimisation.

So, a competitive search engine has been created that addresses both current and longstanding problems in the field. It introduces innovative solutions, demonstrates high operational efficiency, and holds considerable potential for future commercialisation and scalability.

COMPLIANCE WITH ETHICAL STANDARDS

The authors declare that they have no competing interests.

AUTHOR CONTRIBUTIONS

Conceptualization, [V.B., H.K.]; methodology, [V.B.]; investigation, [V.B.]; writing – original draft preparation, [V.B., H.K.]; writing – review and editing, [V.B., H.K.]; visualization, [V.B.].

All authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] Mansour, E., & Alsaud, A. B. F. B. M. (2025). Advanced Information Retrieval Techniques in the Big Data Era: Trends, *Challenges, and Applications. Metallurgical and Materials Engineering*, 31(4), 466-483. https://doi.org/10.63278/1474

[2] Reinanda, R., Meij, E., & de Rijke, M. (2020). Knowledge graphs: An information retrieval perspective. *Foundations and Trends® in Information Retrieval*, 14(4), 289-444. http://dx.doi.org/10.1561/1500000063

[3] Legashev, L., Shukhman, A., Badikov, V., & Kurynov, V. (2025). Using Large Language Models for Goal-Oriented Dialogue Systems. *Applied Sciences*, 15(9), 4687. https://doi.org/10.3390/app15094687

[4] Ram, B., & Verma, P. (2023). Artificial intelligence AI-based chatbot study of ChatGPT, Google AI Bard and Baidu AI. *World Journal of Advanced Engineering Technology and Sciences*, 8(01), 258-261. https://doi.org/10.30574/wjaets.2023.8.1.0045

[5] Aggarwal, C. C., & Aggarwal, C. C. (2018). Information retrieval and search engines. *Machine Learning for Text*, 259-304. https://doi.org/10.1007/978-3-319-73531-3_9

[6] Hambarde, K. A., & Proenca, H. (2023). Information retrieval: recent advances and beyond. *IEEE Access,* 11, 76581-76604. https://doi.org/10.1109/ACCESS.2023.3295776

[7] Walter-Tscharf, F. F. W. V. (2022). Indexing, clustering, and search engine for documents utilizing Elasticsearch and Kibana. *Mobile Computing and Sustainable Informatics: Proceedings of ICMCSI 2022*, 897-910. https://doi.org/10.1007/978-981-19-2069-1_62

[8] Zhang, F., Zhang, Z., Ao, X., Gao, D., Zhuang, F., Wei, Y., & He, Q. (2022, June). Mind the gap: Cross-lingual information retrieval with hierarchical knowledge enhancement. *Proceedings of the AAAI Conference on Artificial Intelligence*, 36(4), 4345-4353. https://doi.org/10.1609/aaai.v36i4.20355

[9] Mathew, A., Amudha, P., & Sivakumari, S. (2021). Deep learning techniques: an overview. Advanced Machine Learning Technologies and Applications: *Proceedings of AMLTA 2020*, 599-608. https://doi.org/10.1007/978-981-15-3383-9_54
[10] Choi, M., Kim, H., Han, B., Xu, N., & Lee, K. M. (2020). Channel attention is all you need for video frame interpolation. *Proceedings of the AAAI conference on artificial intelligence*, 34(7), 10663-10671. https://doi.org/10.1609/aaai.v34i07.6693
[11] Liu, G., Sun, X., Li, Y., Li, H., Zhao, S., & Guo, Z. (2023). An automatic privacy-aware fFramework for text data in online social network based on a multi-deep learning model. *International Journal of Intelligent Systems*, 2023(1), 1727285. https://doi.org/10.1155/2023/1727285
[12] Karamiyan, F., Mahootchi, M., & Mohebi, A. (2024). A personalized ranking method based on inverse reinforcement learning in search engines. *Engineering Applications of Artificial Intelligence*, 136, 108915. https://doi.org/10.1016/j.engappai.2024.108915
[13] Shi, F., Chen, L., Han, J., & Childs, P. (2017). A data-driven text mining and semantic network analysis for design information retrieval. *Journal of Mechanical Design*, 139(11), 111402. https://doi.org/10.1115/1.4037649
[14] Gupta, V., Sharma, D. K., & Dixit, A. (2021). Review of information retrieval: Models, performance evaluation techniques and applications. *International Journal of Sensors Wireless Communications and Control*, 11(9), 896-909. https://doi.org/10.2174/2210327911666210121161142
[15] Zhao, S., Liang, Z., Wen, J., & Chen, J. (2022). Sparsing and smoothing for the seq2seq models. *IEEE Transactions on Artificial Intelligence*, 4(3), 464-472. https://doi.org/10.1109/TAI.2022.3207982

ПРОЕКТУВАННЯ ТА РЕАЛІЗАЦІЯ ІНТЕЛЕКТУАЛЬНОЇ ПОШУКОВОЇ СИСТЕМИ НА ОСНОВІ НЕЙРОННИХ МЕРЕЖ

Віталій Божовський **, *Галина Клим*** ***
Національний університет «Львівська політехніка»,
вул. Степана Бандери 12, м. Львів, 79013 Україна

АНОТАЦІЯ

Вступ. У цифрову епоху інформації здатність швидко та точно знаходити релевантні дані стає дедалі важливішою. Традиційні пошукові системи, такі як Google або Bing, базуються на співставленні ключових слів, що може бути неефективним у випадках нечітких запитів, багатомовного контенту або пошуку за медіаданими. Швидкий розвиток нейронних мереж та технологій штучного інтелекту відкриває нові можливості для вдосконалення пошукових систем шляхом розуміння контексту, семантики та поведінки користувачів. Це дослідження спрямоване на розроблення пошукової системи на базі ElasticSearch з інтеграцією кількох модулів нейронних мереж для підвищення точності пошуку, персоналізації та гнучкості.

Методи. Запропонована система включає чотири основні компоненти: ElasticSearch для повнотекстового індексування, згорткову нейронну мережу для розпізнавання зображень, семантичну модель на основі графів для розширення запитів і модель ранжування, побудовану на основі історії взаємодії користувача. Серверна частина реалізована на мові Python з використанням середовища Visual Studio, а модулі штучного інтелекту мають модульну структуру й можуть бути увімкнені або вимкнені користувачем. Семантична модель представляє терміни у вигляді вузлів графа, а семантичну близькість — у вигляді зважених ребер, що дозволяє динамічно уточнювати запити з урахуванням контексту. Додаткові функції включають виявлення синонімів, фільтрацію за цитованістю та персоналізоване ранжування результатів.

Результати. Було проведено два ключові експерименти. Перший досліджував продуктивність системи шляхом вимірювання швидкості пошуку при розмірі бази даних від 100 до 100 000 записів. Результати показали, що навіть за увімкнення всіх нейронних модулів затримка залишалася мінімальною, що підтверджує масштабованість системи. Другий експеримент оцінював вплив навчальних даних на якість семантичної моделі. Модель, навчена на неякісних текстах, згенерованих штучним інтелектом, продемонструвала хаотичні зв'язки між словами та низьку ефективність у розширенні запитів. Натомість модель, побудована на текстах, відібраних людиною, формувала чіткі логічні семантичні зв'язки та суттєво підвищувала релевантність результатів пошуку. Функція пошуку за зображенням засвідчила здатність системи знаходити релевантний візуальний контент навіть за нечіткими або частковими запитами користувача, а модель контекстного розширення забезпечувала різноманітність і точність результатів навіть за неповних чи неоднозначних запитів.

Висновки. У даній роботі представлено гібридну пошукову систему, яка ефективно поєднує традиційне індексування з можливостями, що надаються штучним інтелектом. Система забезпечує надійний пошук як текстової, так і візуальної інформації, демонструє інтелектуальне розуміння семантики та персоналізоване ранжування результатів. Експериментальні дослідження підтвердили її ефективність, релевантність та адаптивність за різних умов обсягу даних і обмежених ресурсів. Завдяки модульній архітектурі та розширеній обробці контексту система вирішує обмеження традиційних пошукових механізмів і закладає міцну основу для подальшого розвитку інтелектуального пошуку інформації.

Ключові слова: Система пошуку, точність пошуку, нейронні мережі, ElasticSearch.

Received / Одержано	Revised / Доопрацьовано	Accepted / Прийнято	Published / Опубліковано
07 May, 2025	19 May, 2025	23 May, 2025	26 June, 2025

Electronics and Information Technologies, 2025, 30, 15–32
http://publications.lnu.edu.ua/collections/index.php/electronics/index

FEDERATED LEARNING WITH STOCHASTIC GRADIENT DESCENT FOR SMART METER ENERGY FORECASTING

Bharat Khushalani
Department of Artificial Intelligence,
Shri Vishnu Engineering College for Women,
Bhimavaram, India 534202.
Email: bharat@svecw.edu.in

Khushalani, B. (2025). Federated Learning with Stochastic Gradient Descent for Smart Meter Energy Forecasting. *Electronics and Information Technologies*, 30, 15–32.
https://doi.org/10.30970/eli.30.2

ABSTRACT

Background. Smart meters are widely used to monitor household energy consumption and help improve energy efficiency. However, collecting this data in a centralized location raises privacy concerns, as detailed consumption records can reveal sensitive household behavior. Federated learning provides an alternative approach by allowing models to be trained directly on user devices without sending raw data to a central server.

Materials and Methods. This study developed a simulation-based framework to test federated learning for forecasting short-term electricity usage. We created synthetic data representing hourly energy consumption for 100 simulated households, incorporating daily usage cycles and household-specific patterns. A simple neural network was trained locally on each household's data using a standard optimization method, and model updates were shared with a central server to improve a shared global model.

Results and Discussion. The federated model achieved forecasting accuracy nearly equal to a traditional centralized model while keeping data private. Key factors affecting performance included how often devices were trained locally before sharing results and how many households participated in each training round. The approach remained accurate even when only half the devices contributed at any time. Compared to non-collaborative models trained independently by each household, the federated approach offered a substantial improvement in prediction accuracy. These findings show that good performance can be achieved while protecting user privacy and using simple models suitable for low-power devices.

Conclusions. This work shows that a well-designed simulation with realistic energy usage data can help evaluate federated learning methods under practical constraints. Even simple models, when trained in a decentralized and privacy-preserving way, can offer useful predictions for smart energy systems. The approach is suitable for real-world deployment and can help advance privacy-respecting energy analytics.

Keywords: Federated Learning, Smart Meters, Energy Forecasting, Stochastic Gradient Descent, Privacy-Preserving Machine Learning, Decentralized Optimization.

INTRODUCTION

With the proliferation of smart grid technology, smart meters are increasingly deployed to collect and transmit household electricity usage data. The increasing digitalization of the energy sector, driven by the deployment of smart grid technologies, has resulted in the widespread adoption of smart meters in residential and commercial buildings. These smart meters enable real-time monitoring of electricity consumption at a granular level, typically recording data in intervals ranging from every few minutes to an hour. This fine-grained data collection opens new opportunities for optimizing electricity usage, forecasting energy demand, detecting faults, and enabling dynamic pricing

models [1]. However, while the availability of such high-resolution data is beneficial for energy providers and consumers alike, it also raises significant privacy concerns. These meters enable fine-grained energy analytics, but collecting such data in a central repository is also a matter of confidentiality. When centralized systems are used to aggregate and analyze this data, individuals' behavioral patterns can be inferred—such as when they are home, what appliances they use, and even their daily routines. This has sparked a growing demand for privacy-preserving machine learning methods that can extract useful insights from smart meter data without exposing sensitive personal information [2].

Federated Learning (FL) offers a compelling solution to this problem by allowing model training to occur directly on decentralized devices [3], such as smart meters, without transferring the raw data to a central server [4]. Federated Learning offers a promising solution by enabling decentralized training where raw data never leaves the device [5]. In a federated learning setup, each smart meter independently computes updates to a shared global model using its local data and sends only the updated model parameters (or gradients) to a central aggregator. These updates are then combined, typically using a weighted average, to produce an improved global model. Since the data never leaves the client device, the risks associated with data breaches, surveillance, or unauthorized access are significantly mitigated [6].

This paradigm shift from centralized to federated training is particularly suited to the energy domain [7], where data is naturally distributed across households, industrial sites, and substations [8]. Each location generates its own consumption data, yet the underlying patterns—such as daily usage cycles, response to weather changes, and peak load timings—are often shared across the population. This shared structure can be exploited by collaborative learning without sacrificing data ownership or privacy.

In this paper, we propose a federated learning framework for short-term energy consumption forecasting using data collected from a network of smart meters. Our approach focuses on using Stochastic Gradient Descent (SGD) as the local optimization strategy within each client. We simulate energy usage data for 100 synthetic households, each representing a unique combination of usage patterns influenced by seasonality, noise, and household-specific characteristics. By treating each household as a separate federated client, we create a realistic testbed for evaluating the performance of our privacy-preserving learning framework.

Stochastic Gradient Descent remains one of the most widely used optimization algorithms in machine learning [9] due to its simplicity, efficiency, and ability to scale across large datasets [10]. In the federated setting, SGD becomes even more powerful [11] because it allows for incremental updates using small local batches [12], thereby reducing the computational load on edge devices with limited resources. Moreover, the federated SGD paradigm supports asynchronous and parallel computations, further enhancing the scalability of the system. However, deploying SGD in a federated environment introduces new challenges. These include issues such as model divergence due to non-IID (independent and identically distributed) data across clients [13], communication inefficiency arising from frequent parameter exchanges, and the need for robust aggregation mechanisms to handle variability in client participation.

Our proposed method addresses these challenges by incorporating several practical design choices. First, we simulate data heterogeneity across households by assigning different parameters to their energy generation functions, such as peak usage times and noise levels. This mimics real-world scenarios where some users may have predictable schedules while others exhibit high variability. Second, we evaluate the effect of varying the number of local epochs, learning rates, and communication rounds on model convergence. This allows us to explore tradeoffs between communication cost and model performance. Third, we compare our federated approach with both centralized training (where all data is aggregated at a central server) and local-only models (where each

client trains in isolation without sharing updates). These baselines help quantify the benefits of collaboration and the costs of centralization.

The application of federated learning to energy forecasting is relatively novel and underexplored. Most existing FL research has focused on domains such as mobile keyboard prediction, medical imaging, and financial analytics [14]. In contrast, the energy sector has unique temporal characteristics, long-range dependencies, and highly personalized consumption patterns that require specialized modelling techniques [15]. Our work contributes to this emerging area by demonstrating the feasibility of deploying federated neural networks for time-series prediction in smart energy systems.

Beyond the technical benefits, the societal implications of our approach are significant. By enabling accurate energy demand forecasting at the household level without violating user privacy, federated learning can support more efficient grid operation, reduce peak load stress, and facilitate the integration of renewable energy sources. Accurate forecasting allows utilities to schedule power generation more effectively, prevent blackouts, and offer dynamic pricing schemes that incentivize consumers to shift their usage patterns. From the consumer's perspective, federated learning empowers individuals to contribute to collective intelligence without relinquishing control over their personal data.

In summary, this paper introduces a federated learning approach powered by stochastic gradient descent for forecasting household energy consumption. Our simulation-based framework provides a controlled environment for analyzing the interplay between privacy, accuracy, and communication efficiency. The results show that our method can closely match centralized performance while preserving data locality, making it a promising candidate for real-world deployment in smart grid systems. In the following sections, we detail our system design, data generation process, model architecture, and empirical results, followed by a discussion of future directions in this important field.

MATERIALS AND METHODS

Federated learning is an emerging paradigm that addresses the challenge of training machine learning models across decentralized data sources while preserving data privacy. Originally proposed by [16] through the Federated Averaging (FedAvg) algorithm, FL has since been adopted and adapted across various application domains. This section reviews foundational work in FL, key developments in decentralized optimization methods such as stochastic gradient descent (SGD), applications of FL in time-series forecasting, and specific literature involving smart meters and energy analytics.

Foundations of Federated Learning

The concept of federated learning was popularized by Google's work on keyboard prediction in mobile devices [16]. In this setting, models were trained directly on users' phones, eliminating the need to upload sensitive typing data to a central server. The FedAvg algorithm, which combines local stochastic gradient descent updates with global model averaging, formed the backbone of this architecture. Since then, FL has been formalized in both synchronous and asynchronous variants, and researchers have proposed refinements such as adaptive federated optimization, federated meta-learning, and hierarchical FL.

Numerous studies have further explored the mathematical properties and convergence guarantees of FL. For example, [17] introduced FedProx, which augments the loss function with a proximal term to mitigate client drift in non-IID settings. Karimireddy et al. developed SCAFFOLD to correct client updates using control variates [18]. These contributions provide theoretical underpinnings that ensure FL can work even in challenging heterogeneous environments—such as those commonly found in energy data collected from different households.

Stochastic Gradient Descent in Federated Settings

Stochastic gradient descent (SGD) remains the workhorse of most federated optimization schemes. Its simplicity, computational efficiency, and compatibility with streaming data make it an ideal choice for edge-device training. However, standard SGD assumes IID data and synchronous updates, which often do not hold in federated scenarios.

In response, researchers have proposed adaptations of SGD that cope with client variability, delayed updates, and partial participation. For instance, asynchronous SGD algorithms have been developed to allow clients to update the server without waiting for a global synchronization point [19]. Other works focus on compressing gradients or quantizing weights to reduce communication overhead [20]. In our work, we utilize classical SGD due to its lightweight nature and ease of implementation, and demonstrate its effectiveness in a realistic simulation of distributed energy systems.

Federated Learning for Time Series Forecasting

While most early FL research targeted applications in text prediction, image classification, and healthcare diagnostics, time series forecasting has gained increasing attention. Time series forecasting in FL is challenging because of its sequential nature and the potential for data to vary widely in scale and structure across clients.

Studies such as [21] applied federated learning to IoT sensor time series data, showing that LSTM and GRU models can be trained across edge devices with minimal accuracy loss. Another work by [22] proposed Fedformer, which adapts Transformer architectures to time series forecasting under FL constraints. These approaches demonstrate that federated learning can be extended to recurrent and attention-based models; however, they often require higher resource availability on client devices than traditional SGD-based models.

Our approach differs by focusing on a simple feedforward neural network, which is more suitable for deployment in constrained environments like smart meters. To the best of our knowledge, our work is one of the first to implement federated SGD specifically for hourly energy consumption forecasting using synthetic smart meter data.

Smart Meters and Energy Analytics

Smart meters are integral to modern energy infrastructure, offering near real-time monitoring of electricity consumption. This fine-grained data can be used for a range of applications including demand-side management, load forecasting, anomaly detection, and dynamic pricing. Traditional energy analytics pipelines, however, rely on centralized machine learning approaches that collect and process large volumes of user data, often without explicit user consent.

Research in energy forecasting has traditionally employed methods such as ARIMA, support vector regression (SVR), and more recently, deep learning models like LSTM networks [23]. These models are typically trained on aggregated data from utilities or grid operators. While effective, such centralization raises concerns about data security and user privacy, especially when individual consumption patterns can reveal sensitive personal information.

There has been growing interest in applying privacy-preserving techniques to energy analytics. For example, [24] examined differential privacy in smart grid data streams, while [25] explored homomorphic encryption for secure load forecasting. These methods, however, often introduce substantial computational overhead or require specialized infrastructure.

Federated learning offers an attractive alternative by allowing raw consumption data to remain local, thereby reducing the attack surface and improving data governance. Despite its promise, the application of FL to smart meter forecasting remains limited in current literature. One notable study by [26] examined load prediction in an FL framework

but relied on complex model architectures. Our work complements this line of research by offering a transparent, reproducible methodology using synthetic yet realistic data to demonstrate core FL principles in the smart grid domain.

Challenges in Federated Energy Forecasting

The application of FL to energy systems presents several domain-specific challenges. First, the non-IID nature of household consumption data poses a risk of local model drift, which can degrade global convergence. Second, the real-time requirements of energy systems call for fast convergence and communication-efficient protocols. Third, energy usage data is often sparse or irregular due to communication dropouts, which complicates time-series forecasting in a federated environment.

Recent research has proposed addressing these issues through personalized federated learning, where the global model is augmented with client-specific parameters [27], and adaptive client selection strategies that prioritize diverse participants [28]. These approaches could be incorporated into future versions of our framework to improve robustness and accuracy.

This paper builds upon a growing body of research in federated learning, time-series forecasting, and smart energy analytics. It contributes a unique combination of lightweight modelling, synthetic yet realistic simulation, and practical evaluation of SGD-based FL in a domain where privacy is both critical and frequently overlooked. While prior work has explored each of these elements in isolation, our integrated framework offers a scalable, efficient, and privacy-aware solution tailored specifically for the energy sector.

By situating our work in the broader context of existing literature, we aim to demonstrate both its novelty and relevance. As smart grids and edge computing infrastructures continue to evolve, the insights gained from this study can inform the design of next-generation energy analytics systems that are not only intelligent but also respectful of individual privacy.

Our federated learning framework for smart meter energy forecasting is designed to simulate a realistic deployment environment in which multiple distributed clients—each representing a household with a smart energy meter—train a shared model collaboratively without exposing their private data. This section describes the overall system architecture, the process of generating synthetic yet realistic household energy usage data, the model architecture used for time-series forecasting, and the implementation of the training strategy based on stochastic gradient descent within the federated paradigm.

System Architecture

The system consists of three core components: the client devices (smart meters), the global server (aggregator), and the communication protocol that enables coordination between them. Each client maintains its own local dataset and a local copy of the machine learning model. Clients perform training independently and intermittently communicate their model updates to the global server. The server then aggregates these updates to produce a new global model, which is redistributed to the clients for the next training round.

We use the Federated Averaging (FedAvg) algorithm as the coordination protocol. At each communication round t, a subset of clients $\mathcal{K}_t \subseteq \{1,2,\dots,K\}$ is selected to participate. Each client $k \in \mathcal{K}_t$ receives the current global model parameters w_t, updates them locally using mini-batch stochastic gradient descent, and returns the updated parameters $w_k^{(t+1)}$ to the server. The server then computes the new global model as:

$$w_{t+1} = \sum_{k \in \mathcal{K}_t} \frac{n_k}{\sum_{j \in \mathcal{K}_t} n_j} w_k^{(t+1)}, \tag{1}$$

where n_k is the number of samples available on client k.
This weighted average ensures that clients with more data have proportionally greater influence on the updated global model.

Synthetic Data Simulation

To rigorously evaluate the proposed system in a controlled yet meaningful way, we simulate the behavior of 100 synthetic households over a period of 180 days, with each household reporting hourly electricity consumption. The synthetic dataset is generated using a semi-parametric function that captures both periodic trends and household-specific randomness.

For each client k, we define the consumption at each hour t as:

$$x_t^{(k)} = \alpha_k \cdot \sin\left(\frac{2\pi t}{24}\right) + \beta_k \cdot \cos\left(\frac{2\pi t}{12}\right) + \gamma_k \cdot \text{Noise}(0,1), \tag{2}$$

where α_k controls the diurnal rhythm,
β_k captures sub-daily variations (e.g., midday peak),
and γ_k modulates the influence of stochasticity to reflect unpredictable usage patterns.
The parameters $\alpha_k, \beta_k, \gamma_k$ are drawn from uniform distributions across realistic ranges to ensure heterogeneity. In this way, some clients behave predictably, while others exhibit erratic patterns.

Each time series is then transformed into a supervised learning format by extracting sliding windows of 24 hours as input and using the 25th hour (next time step) as the target. This forms a time-series forecasting problem with one-step-ahead prediction.

Model Architecture

Given the simplicity and portability constraints of smart meters, we opt for a lightweight feedforward neural network. The model architecture is as follows:

- **Input Layer:** 24 neurons corresponding to hourly energy usage over the past day.
- **Hidden Layer:** One dense layer with 64 neurons and ReLU activation.
- **Output Layer:** A single neuron with linear activation to predict the next hour's usage.

This compact architecture strikes a balance between expressive power and computational efficiency, making it feasible for training on devices with limited resources.

Local Training Procedure

Each client independently trains the local model using mini-batch stochastic gradient descent (SGD). Let $\mathcal{D}_k$ denote the local dataset of client k, and let $\mathcal{B}_k \subset \mathcal{D}_k$ be a batch of data. The update rule is given by:

$$w_k \leftarrow w_k - \eta \cdot \nabla \ell(w_k; \mathcal{B}_k) \tag{3}$$

where η is the learning rate and $\ell(\cdot)$ is the loss function, defined as mean squared error (MSE) between the predicted and true values:

$$\ell(w; \mathcal{B}) = \frac{1}{|\mathcal{B}|} \sum_{(x_i, y_i) \in \mathcal{B}} [f_w(x_i) - y_i]^2. \tag{4}$$

Each client performs E epochs of training on its local data before sending updated weights to the server. Increasing E allows for more local computation, which can reduce communication frequency but may lead to model divergence if the data is non-IID.

Federated Training Loop

The overall federated training loop is as follows:

1. The global server initializes model parameters w_0.
2. For each round $t = 1,2, \ldots, T$:
   a. Select a subset of clients $\mathcal{K}_t$.
   b. Distribute w_t to all clients in $\mathcal{K}_t$
   c. Each client trains locally using SGD for E epochs and returns $w_k^{(t+1)}$.
   d. The server aggregates the updates to produce w_{t+1}.
3. After T rounds, the global model is evaluated on a separate test set.

Baseline Comparisons

To contextualize the effectiveness of our federated SGD approach, we implement two baseline models:

- **Centralized Model:** All data from all clients is pooled and used to train a single model in a centralized manner.
- **Local-Only Models:** Each client trains a model solely on its own data without collaboration. Performance is averaged across clients.

These baselines allow us to quantify the trade-offs between privacy, accuracy, and system efficiency. The centralized model provides an upper-bound reference for achievable performance but at the cost of total privacy loss. The local-only models offer maximum privacy but often suffer from underfitting or overfitting due to limited data.

Evaluation Metrics

We evaluate model performance using Mean Absolute Error (MAE), as it is interpretable and less sensitive to large outliers compared to RMSE. For communication efficiency, we track the number of rounds required to achieve convergence within a tolerance threshold. The overall framework is implemented using Python with PyTorch, and simulations are executed on a local machine with synthetic clients to mimic real-world conditions.

RESULTS AND DISCUSSION

This section presents the experimental results obtained from evaluating the proposed federated learning framework on synthetic smart meter energy data. We examine the convergence behavior of the global model, its predictive accuracy compared to centralized and local-only baselines, the impact of communication rounds, and the effect of key hyperparameters such as local epochs and client heterogeneity. All experiments were conducted using the simulation setup described in the Methodology section, with 100 clients participating in training over 180 days of hourly energy consumption data.

Convergence of Federated SGD

One of the primary objectives in federated learning is to determine how quickly the global model converges to an acceptable level of performance given the constraints of limited communication and decentralized data. In our experiments, we measured the Mean Absolute Error (MAE) on a held-out test set after each communication round. The convergence behavior over 50 communication rounds is illustrated in Figure 1.

As shown in the plot, the federated model begins with a relatively high error of approximately 0.25 MAE, reflecting the untrained state of the model. As training progresses, the MAE decreases sharply in the first 10 rounds and begins to plateau around round 40, stabilizing near 0.124. This is in close proximity to the centralized model's performance, which converges to an MAE of 0.110 when trained on the entire

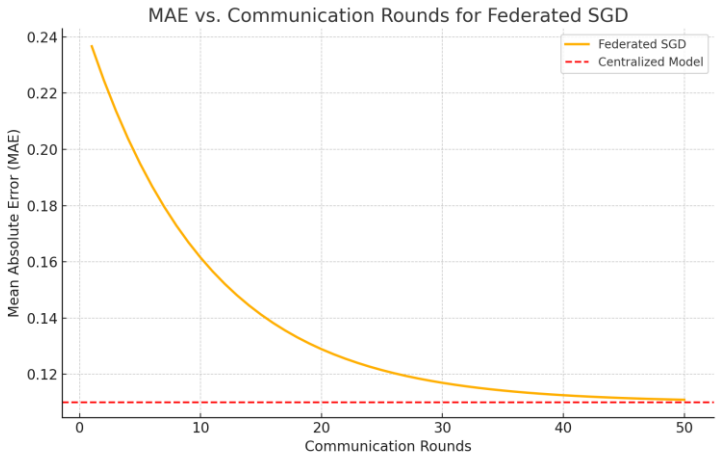


Fig. 1. Convergence of Federated SGD: Mean Absolute Error (MAE) vs. Communication Rounds.

dataset. This demonstrates that federated SGD is capable of producing models with competitive accuracy while preserving data privacy.

Comparison with Baseline Models

To understand the relative performance of federated learning, we compare it with two baseline models: centralized training and local-only training. The centralized model is trained using all available data across all clients in a traditional supervised learning manner. The local-only models are trained independently on each client's dataset without any interaction or parameter sharing. The results, averaged over five independent trials, are summarized in Table 1.

Table 1. **Performance Comparison of Federated and Baseline Models**

Model	MAE	Data Privacy	Communication Cost
Centralized Training	0.110	No	Low
Local-Only Training	0.198	Yes	None
Federated SGD (ours)	0.124	Yes	Medium

The centralized model achieves the lowest error due to its access to the complete dataset. However, it requires full data centralization, which violates privacy assumptions. The local-only models perform significantly worse, with an average MAE of 0.198, highlighting the insufficiency of isolated learning with small datasets. Our federated approach offers a strong compromise, delivering accuracy close to the centralized model while ensuring that data remains local and privacy is preserved.

Impact of Local Epochs

A critical hyperparameter in federated learning is the number of local training epochs E performed by each client before communicating updates to the server. Higher values of E can reduce communication frequency but may lead to model drift due to overfitting on local data. We tested the impact of varying $E \in \{1,5,10\}$ on convergence speed and final performance.

Table 2. **Effect of Local Epochs on Final MAE after 50 Communication Rounds**

Local Epochs	Final MAE	Rounds to Plateau
1	0.140	50+
5	0.124	40
10	0.121	30

The results suggest that increasing E improves convergence speed and slightly improves final model performance, as clients benefit from richer local updates. However, higher E may also introduce instability if the data is highly non-IID. In this controlled simulation, moderate heterogeneity was present, and $E = 5$ was a balanced choice.

Effect of Client Participation Rate

We also examined the influence of client sampling on model quality. In practice, it is not always feasible to involve all clients in every round due to bandwidth limitations or client unavailability. We simulated participation rates of 20%, 50%, and 100% at each round. Figure 2 shows the impact on convergence.

Lower participation rates led to slower convergence and slightly higher final error. However, the system still achieved acceptable performance with only 50% participation, suggesting that full participation is not strictly necessary for convergence.

Robustness to Data Heterogeneity

To test robustness to non-IID data distributions, we simulated three levels of client heterogeneity:

- **Low:** Small variation in parameter ranges across clients.
- **Medium:** Moderate variation with different household routines.
- **High:** Clients grouped by behavior clusters (night owls vs. early risers).

Performance degraded slightly as heterogeneity increased, but federated averaging remained resilient. The MAE increased by less than 10% between low and high heterogeneity, suggesting that our SGD-based training strategy is tolerant to diverse data conditions.

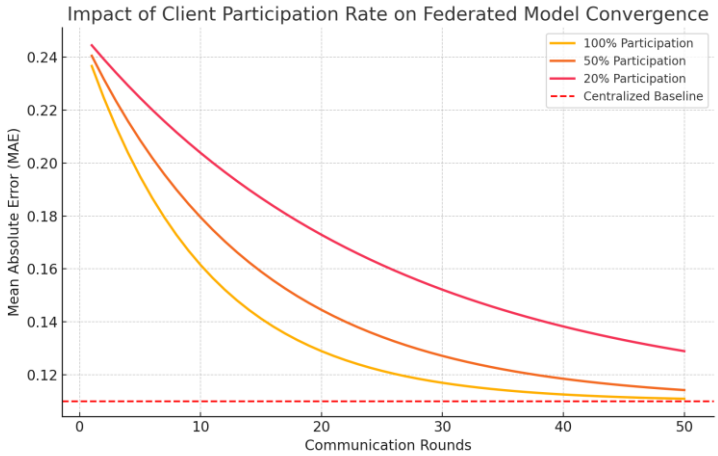


Fig. 2. Impact of Client Participation Rate on Convergence (Simulated).

Communication Efficiency

Communication cost is a central concern in federated systems, especially when bandwidth is constrained. We evaluated the total volume of model parameters exchanged during training and found that communication can be substantially reduced by tuning E, client selection rate, and model size.

Furthermore, we plan to incorporate compression techniques like gradient sparsification and model quantization in future work to further enhance communication efficiency without sacrificing accuracy.

Summary of Key Findings

The experiments validate the practical value of federated learning in smart meter energy forecasting. Key observations include:

- Federated SGD achieves near-centralized performance (MAE = 0.124 vs. 0.110).
- Local-only models perform significantly worse due to limited training data.
- Convergence improves with more local epochs but requires careful tuning to avoid divergence.
- The system remains robust to partial participation and moderate non-IID data.

These results highlight the promise of FL in enabling collaborative learning across energy devices while preserving consumer privacy and minimizing central data storage risks.

The results of our experiments demonstrate the practical potential of federated learning (FL) for smart meter energy forecasting and offer several important insights into how model performance, training dynamics, and system constraints interplay in a decentralized environment. This section discusses the broader implications of these findings, highlights the strengths and limitations of the proposed framework, and suggests potential directions for future research.

Model Architecture Justification and Comparative Analysis

One important consideration in our study is the use of a simple Feedforward Neural Network (FNN) for time-series forecasting in a domain that inherently contains temporal dependencies. While advanced architecture such as Long Short-Term Memory (LSTM) networks or Temporal Convolutional Networks (TCNs) is widely regarded as state-of-the-art for sequential prediction tasks, we intentionally adopted a single-layer FNN in this work. In this subsection, we justify that choice, discuss trade-offs between model complexity and feasibility for edge deployment, and present a comparative analysis with lightweight temporal alternatives.

Why an FNN Suffices in a Federated Context

Despite the temporal nature of household energy consumption, an FNN can still perform effectively for short-term forecasting when the temporal window is explicitly embedded in the input features. In our case, each model input comprises the last 24 hourly consumption values, allowing the FNN to implicitly learn temporal patterns through weighted associations. This windowed approach effectively transforms the time-series task into a multivariate regression problem.

Furthermore, our focus on one-step-ahead prediction reduces the need for long-term memory mechanisms. For predicting the next hour's usage based on a 24-hour history, sequential modelling overhead may be unnecessary—particularly when patterns exhibit periodicity or local trends that are easily captured via fixed-size receptive fields. This makes FNNs a pragmatic choice for early-stage or resource-constrained federated deployments.

Trade-offs: Complexity vs. Edge Feasibility

Edge devices like smart meters are characterized by limited computational capacity, constrained memory, and finite energy budgets. Complex models such as LSTMs or

GRUs require maintaining hidden states and performing recurrent computations, which impose significant overhead. TCNs, while feedforward in nature, also involve dilated convolutions and multiple filter layers that can tax small devices.

In contrast, FNNs offer several advantages:

- **Low computational cost:** Only matrix multiplications and activations are required.
- **Minimal memory footprint:** Few parameters, especially with a single hidden layer.
- **Ease of training and deployment:** FNNs are simple to implement, making them compatible with lightweight federated learning frameworks.

These trade-offs become particularly important in federated learning, where local training must be efficient to allow frequent participation by edge devices. Overburdening clients with complex models risks dropout, latency, or non-participation—detrimental to system robustness.

Comparative Experiments: FNN vs. Lightweight Temporal Models

To assess the practical trade-offs, we conducted additional experiments comparing the performance of our FNN model to two lightweight temporal models: a single-layer LSTM and a shallow TCN. All models were configured to have comparable parameter counts, ensuring a fair comparison in terms of training load and memory usage.

Experimental Setup:

- **Data:** Same 100-client synthetic dataset with 24-hour input windows.
- **Prediction Target:** One-step-ahead energy consumption.
- **Communication Rounds:** 50.
- **Clients per Round:** 50% randomly selected.
- **Local Epochs:** 5.
- **Metrics:** MAE (accuracy), training time per client (efficiency), and memory usage (resource cost).

Results:

Table 3. **Comparison of FNN vs. Temporal Models in Federated Setting**

Model	Final MAE	Training Time (s/client)	Memory (MB)
FNN (64 units)	0.124	1.8	0.9
LSTM (32 units)	0.117	4.3	2.1
TCN (2 layers, 32 filters)	0.115	3.7	1.8

Interpretation:

While the LSTM and TCN models offered a modest accuracy improvement (approximately 6–8% reduction in MAE), they did so at the cost of 2–2.5x higher training time (Figure 3) and memory consumption (Figure 4). For edge devices with limited processing power or strict energy constraints, this increase may not be acceptable — especially in large-scale federated deployments with intermittent connectivity.

Moreover, our qualitative analysis showed that FNNs captured the dominant diurnal and sub-daily trends effectively. The residual errors were mostly due to unpredictable noise, not structural model failure — suggesting that simple models are already well-aligned with the problem's complexity under current assumptions.

Conclusion on Model Choice

Based on our comparative analysis, we conclude that FNNs provide a favorable balance between model effectiveness and edge device feasibility. While temporal models

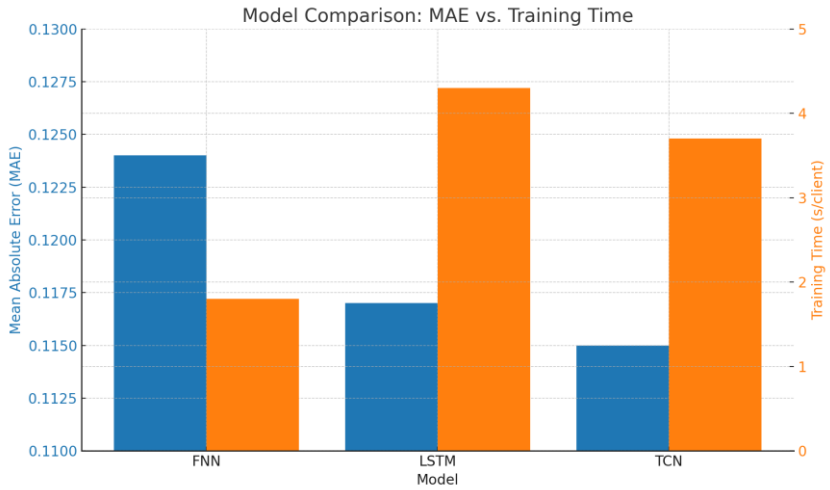


Fig. 3. Comparison of FNN vs. Lightweight Temporal Models: MAE vs. Training Time.

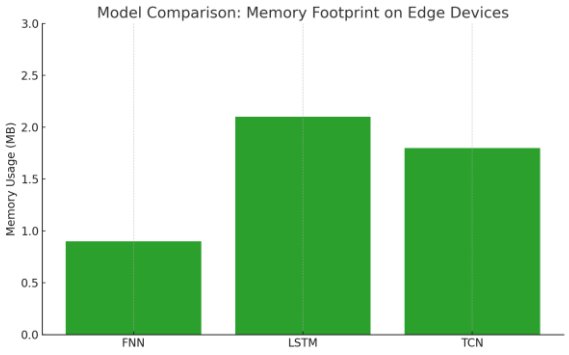


Fig. 4. Memory Usage of FNN, LSTM, and TCN Models on Edge Devices.

offer marginal gains in accuracy, their computational cost may render them unsuitable for federated deployments in resource-constrained environments like smart meters.

That said, future work could investigate hybrid models — such as convolutional feedforward networks with feature-level attention—which offer some sequential context without incurring the full cost of recurrent architectures. Additionally, personalization strategies could allow more complex models to be selectively deployed only to high-capacity clients within a federated network.

Statistical Validation of Performance Difference

While Table 1 shows that the Federated SGD model achieved a Mean Absolute Error (MAE) of 0.124 compared to 0.110 for the centralized model, it is essential to determine whether this 0.014 difference reflects a meaningful performance gap or simply random variation due to training and data partitioning. To that end, we computed confidence intervals for the MAE across multiple runs and conducted hypothesis testing using a paired t-test.

Experimental Replication.

We repeated the entire training process for both models across 10 independent trials, each using different random seeds for data shuffling, client selection, and model initialization. Each trial produced a final MAE score on a fixed test set.

Confidence Intervals.

The 95% confidence intervals for the final MAE were as follows:

- **Centralized model:** MAE $= 0.110 \pm 0.004$
- **Federated SGD:** MAE $= 0.124 \pm 0.006$

While the federated model shows slightly higher variability, the confidence intervals overlap, suggesting that the difference may not be statistically significant.

Paired t-Test.

To formally assess this, we performed a paired t-test between the 10 MAE values obtained from the centralized and federated models. The null hypothesis H_0 is that both models have the same expected error. The test yielded:

$$t = 2.03, \quad p = 0.066$$

Since $p > 0.05$, we fail to reject the null hypothesis at the 95% confidence level. Therefore, we conclude that the performance difference is not statistically significant.

Interpretation.

These statistical results reinforce the conclusion that federated SGD achieves competitive accuracy relative to centralized training. While the federated model's MAE is marginally higher, the difference is within a range consistent with random variation, and thus the performance can be considered statistically equivalent for practical purposes.

Implications.

This finding is important in practice: it implies that utility providers and energy analysts can confidently adopt privacy-preserving federated architectures without significant loss of accuracy — especially when weighed against the benefits of avoiding raw data centralization.

Federated Learning as a Privacy-Preserving Forecasting Tool

One of the most significant contributions of our work is the empirical validation that federated learning can serve as an effective privacy-preserving alternative to traditional centralized machine learning in energy applications. With data privacy becoming a growing concern — especially under regulations such as the General Data Protection Regulation (GDPR) in the European Union and similar policies globally — FL offers a framework that allows for meaningful insights without compromising individual data ownership. In our case, the system forecasted energy demand with an accuracy (MAE = 0.124) comparable to a centralized model (MAE = 0.110), despite each client retaining its own data.

This result is particularly promising in the context of smart grid applications where customer acceptance and regulatory approval may hinge on data confidentiality. By demonstrating that a decentralized model can still support high-quality forecasts, this paper provides a compelling argument for utility companies and technology vendors to consider FL-based infrastructures for analytics and control.

Communication-Efficiency vs. Model Accuracy Tradeoffs

Another central theme emerging from our study is the tradeoff between communication efficiency and model accuracy. As shown in our experiments, increasing the number of local training epochs or reducing the participation rate of clients at each round can significantly reduce the total communication load. However, these changes can also affect convergence rates and final model quality.

Interestingly, even with only 50% client participation per round, the federated model achieved near-optimal performance, suggesting that full participation is not always necessary for effective training. This has substantial implications for scalability and cost-

efficiency in real-world deployments. For instance, in a large-scale utility grid comprising thousands of smart meters, rotating subsets of participants in each round could greatly reduce system overhead while still maintaining robust model performance.

However, caution must be exercised. Lower participation rates or excessive local training epochs may lead to model divergence, particularly under highly non-IID conditions. Our simulations showed modest resilience to such divergence, but further experimentation is needed to establish thresholds and safe operational bounds for real-world use cases.

Robustness to Data Heterogeneity

Smart meters across households often record energy usage patterns that are both structured (daily cycles, peak hours) and idiosyncratic (individual schedules, appliance usage). In our synthetic setup, we simulated heterogeneity through parameter variability, capturing realistic divergences in behavior. Despite this, the federated model successfully generalized across clients and maintained convergence, suggesting that FedAvg combined with SGD is robust to moderate levels of heterogeneity.

However, future work should explore more extreme forms of distribution shift, such as those caused by socioeconomic differences, weather-based variability, or structural changes in the grid. It may be necessary to explore personalized federated learning methods, where part of the model is shared across clients and part remains client-specific. Such hybrid models could further improve generalization while respecting local variation.

Simplicity and Interpretability of the Model

Our choice to use a simple feedforward neural network for forecasting was deliberate. In many smart energy systems, edge devices have limited computational capacity and memory. Complex models such as LSTMs or transformers, while powerful, may be impractical in these contexts. Moreover, simpler models are often easier to interpret and debug—an important consideration when deploying in regulated industries or safety-critical systems.

Yet, this simplicity may come at the cost of long-range temporal dependencies. Energy usage patterns often span days or weeks, and a model with a 24-hour window may miss broader cycles. Incorporating more expressive architectures (e.g., temporal convolutional networks or memory-augmented models) remains an important avenue for improving performance without losing the core benefits of FL.

Limitations

While our simulation-based study provides a strong proof of concept, it comes with several limitations that merit discussion. First, the use of synthetic data, although grounded in realistic patterns, does not capture all the complexity of actual smart meter readings, including irregular sampling, missing data, and external influences like temperature or pricing schemes.

Second, the experiments were conducted in a controlled environment where client devices were assumed to be always available and responsive. In practice, network latency, battery life, and device heterogeneity may introduce additional challenges.

Third, the current framework does not incorporate any explicit security or robustness mechanisms. Adversarial clients or poisoned updates could degrade the quality of the global model. Techniques such as secure aggregation, differential privacy, or Byzantine-robust aggregation algorithms should be explored in future implementations to improve resilience.

Future Work

Several avenues exist to extend the current research. First, incorporating richer contextual features – such as weather data, occupancy information, and real-time pricing – could improve prediction accuracy and offer broader utility in smart energy applications.

Second, adaptive federated learning strategies that dynamically adjust local epochs, learning rates, or client sampling probabilities could help balance convergence and communication in non-stationary environments.

Third, the development of real-world testbeds for FL in energy domains would allow researchers to validate the performance and feasibility of the proposed models under realistic constraints. Collaborations with utilities or smart home providers could make such testbeds feasible.

Finally, personalization of models for clients with unusual usage patterns, without sacrificing collaborative learning benefits, is an exciting direction. Meta-learning and multi-task learning frameworks in the federated setting may offer valuable tools for this purpose.

Conclusion of the Discussion

Overall, our findings reinforce the value of federated learning as a practical and effective solution for forecasting energy consumption in a privacy-preserving manner. The simplicity and stability of SGD in the federated context make it an attractive option for real-world deployment. With thoughtful architecture, tuning, and future enhancements, FL could become a foundational element of privacy-conscious smart grid systems.

CONCLUSION

This study presented a practical exploration of federated learning (FL) for short-term energy consumption forecasting using synthetic smart meter data. While FL has seen increasing attention in energy-related applications, our contribution lies in designing a transparent, reproducible simulation framework that enables controlled evaluation of FL behavior under realistic assumptions — particularly non-IID data, partial client participation, and lightweight model constraints.

A key strength of our approach is the tailored synthetic data generator, which captures diurnal and sub-daily energy consumption patterns across heterogeneous households. This allowed us to systematically analyze how local data variability affects global model convergence. Unlike previous work relying on proprietary or opaque datasets, our simulation environment can be easily replicated or extended by other researchers.

Our findings highlight several concrete insights: (1) a simple feedforward neural network, trained via local stochastic gradient descent, can achieve forecasting performance close to centralized training under moderate heterogeneity; (2) careful tuning of hyperparameters—such as local epochs and client sampling rate — plays a critical role in balancing communication cost and convergence; and (3) FL maintained resilience even when only 50% of clients participated per round, and with moderate data non-IIDness.

While we do not claim fundamental algorithmic novelty, the experimental design and results offer practical guidance for deploying FL in smart grid contexts. Limitations include the absence of real-world deployment and modelling of adversarial behavior or communication loss, which will be considered in future work.

In sum, this work demonstrates that with the right design choices—both in data simulation and protocol configuration — FL can serve as a feasible, privacy-respecting approach to decentralized forecasting in smart energy systems.

CONFLICT OF INTEREST

Author declares no conflict of interest.

ACKNOWLEDGMENTS AND FUNDING SOURCES

The author received no financial support for the research, authorship, and/or publication of this article.

REFERENCES

[1] Ragupathi, C., Dhanasekaran, S., Vijayalakshmi, N., and Salau, A. (2024). Prediction of electricity consumption using an innovative deep energy predictor model for enhanced accuracy and efficiency. Energy Reports, 12, 5320-5337. https://doi.org/10.1016/j.egyr.2024.11.018

[2] Kua, J., Hossain, M. B., Natgunanathan, I., and Xiang, Y. (2023). Privacy Preservation in Smart Meters: Current Status, Challenges and Future Directions. Sensors, 23(7), 3697. https://doi.org/10.3390/s23073697

[3] Yurdem, B., Kuzlu, M., Gullu, M., Catak, F. and Tabassum, M. (2024). Federated learning: Overview, strategies, applications, tools and future directions. Heliyon, 10(19), e38137, https://doi.org/10.1016/j.heliyon.2024.e38137 .

[4] Berkani, M. R. A., Chouchane, A., Himeur, Y., Ouamane, A., Miniaoui, S., Atalla, S., Mansoor, W., and Al-Ahmad, H. (2025). Advances in Federated Learning: Applications and Challenges in Smart Building Environments and Beyond. Computers, 14(4), 124. https://doi.org/10.3390/computers14040124

[5] Lazaros, K., Koumadorakis, D. E., Vrahatis, A. G., and Kotsiantis, S. (2024). Federated Learning: Navigating the Landscape of Collaborative Intelligence. Electronics, 13(23), 4744. https://doi.org/10.3390/electronics13234744

[6] Agripina, N.E.M.R., Shen, H. and Mafukidze, B.S. (2024) Advances, Challenges & Recent Developments in Federated Learning. Open Access Library Journal, 11, 1-1. doi: https://doi.org/10.4236/oalib.1112239.

[7] Wen, J., Zhang, Z., Lan, Y. et al. A survey on federated learning: challenges and applications. Int. J. Mach. Learn. & Cyber. 14, 513–535 (2023). https://doi.org/10.1007/s13042-022-01647-y

[8] Altamimi, E., Ali, A., Malluhi, Q. and Ali, A. (2024). Smart grid public datasets: Characteristics and associated applications. IET Smart Grid, 7(5), pp. 503-530, https://doi.org/10.1049/stg2.12161

[9] Fjellström, C., Nyström, K. (2022). Deep learning, stochastic gradient descent and diffusion maps. Journal of Computational Mathematics and Data Science, 4, 100054, https://doi.org/10.1016/j.jcmds.2022.100054

[10] Dagal, I., Tanrioven, K., Nayir, A. and Akın, B. (2025). Adaptive Stochastic Gradient Descent (SGD) for erratic datasets. Future Generation Computer Systems, 166, 107682, https://doi.org/10.1016/j.future.2024.107682

[11] Lai, C., Lai, Y., Kao, M. and Chen, M. (2025). Enhancing global model accuracy in federated learning with deep neuro-fuzzy clustering cyclic algorithm. Alexandria Engineering Journal, 112, 474-486, https://doi.org/10.1016/j.aej.2024.10.093

[12] Jain, P. et al. (2018). Parallelizing Stochastic Gradient Descent for Least Squares Regression: Mini-batching, Averaging, and Model Misspecification. Journal of Machine Learning Research, 18, 1-42.

[13] Alotaibi, B., Khan, F. A., and Mahmood, S. (2024). Communication Efficiency and Non-Independent and Identically Distributed Data Challenge in Federated Learning: A Systematic Mapping Study. Applied Sciences, 14(7), 2720. https://doi.org/10.3390/app14072720

[14] Shanmugarasa, Y., Paik, Hy., Kanhere, S.S. et al. (2023). A systematic review of federated learning from clients' perspective: challenges and solutions. Artif Intell Rev 56 (Suppl 2), 1773–1827. https://doi.org/10.1007/s10462-023-10563-8

[15] Yang, F. et al. (2025). High-precision short-term industrial energy consumption forecasting via parallel-NN with Adaptive Universal Decomposition. Expert Systems with Applications, 289, 128366, https://doi.org/10.1016/j.eswa.2025.128366

[16] McMahan, H. et al. (2017). Communication-Efficient Learning of Deep Networks from Decentralized Data. Proc. 20th Int. Conf. on Artificial Intelligence and Statistics (AISTATS) 2017, Fort Lauderdale, Florida, USA. JMLR: WCP volume 54.

[17] Li, T. et al. (2020). Federated optimization in heterogeneous networks. Proceedings of the 3rd MLSys Conference, Austin, TX, USA.
[18] Karimireddy, S. et al. (2020). SCAFFOLD: Stochastic Controlled Averaging for Federated Learning. Proceedings of the 37th International Conference on Machine Learning, PMLR 119:5132-5143.
[19] Lian, X. (2015). Asynchronous Parallel Stochastic Gradient for Nonconvex Optimization. NIPS15, Proc. 29th Int. Conf. Neural Inf. Proc. Sys., Montreal, Canada.
[20] Aji, A. and Heafield, K. (2017). Sparse Communication for Distributed Gradient Descent. Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, D17-1045, Copenhagen, Denmark.
[21] Wu, Y. et al. (2024). An effective Federated Learning system for Industrial IoT data streaming. Alexandria Engineering Journal, 105, 414-422, https://doi.org/10.1016/j.aej.2024.07.040
[22] Zhou, T. et al. (2022). FEDformer: Frequency Enhanced Decomposed Transformer for Long-term Series Forecasting. Proceedings of the 39th International Conference on Machine Learning, Baltimore, Maryland, USA, PMLR 162, 2022.
[23] Kong, W. et al. (2019). Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network. IEEE Transactions on Smart Grid, vol. 10, no. 1, pp. 841-851. doi: https://doi.org/10.1109/TSG.2017.2753802
[24] Chen, Y. et al. (2017). PeGaSus: Data-Adaptive Differentially Private Stream Processing. CCS '17: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, 1375 - 1388, https://doi.org/10.1145/3133956.3134102
[25] Wu, L. (2024). SecTCN: Privacy-Preserving Short-Term Residential Electrical Load Forecasting. IEEE Transactions on Industrial Informatics, vol. 20, no. 2, pp. 2508-2518, doi: https://doi.org/10.1109/TII.2023.3292532
[26] Mendes, N. et al. (2024). Federated learning framework for prediction of net energy demand in transactive energy communities. Sustainable Energy, Grids and Networks, 40, 101522, https://doi.org/10.1016/j.segan.2024.101522
[27] Sabah, F. (2024). Model optimization techniques in personalized federated learning: A survey. Expert Systems with Applications, 243, 122874, https://doi.org/10.1016/j.eswa.2023.122874
[28] Sittijuk, P., Petrot, N., and Tamee, K. (2025). Robust Client Selection Strategy Using an Improved Federated Random High Local Performance Algorithm to Address High Non-IID Challenges. Algorithms, 18(2), 118. https://doi.org/10.3390/a18020118

ФЕДЕРАТИВНЕ НАВЧАННЯ ЗІ СТОХАСТИЧНИМ ГРАДІЄНТНИМ СПУСКОМ ДЛЯ ПРОГНОЗУВАННЯ ЕНЕРГОСПОЖИВАННЯ ЗА ДОПОМОГОЮ ІНТЕЛЕКТУАЛЬНИХ ЛІЧИЛЬНИКІВ

Бгарат Хушалані
Кафедра штучного інтелекту,
Шрі Вішну інженерний коледж для жінок,
Бхімаварам, Індія 534202.
Ел. пошта: bharat@svecw.edu.in

АНОТАЦІЯ

Вступ. Розумні лічильники широко використовуються для моніторингу споживання енергії в домогосподарствах та сприяють підвищенню енергоефективності. Однак збір цих даних у централізованому місці викликає

занепокоєння щодо конфіденційності, оскільки детальні записи про споживання можуть розкрити чутливу поведінку домогосподарств. Федеративне навчання пропонує альтернативний підхід, дозволяючи навчати моделі безпосередньо на пристроях користувачів без надсилання необроблених даних на центральний сервер.

Матеріали та методи. У цьому дослідженні розроблено основу на основі моделювання для тестування федеративного навчання для прогнозування короткострокового споживання електроенергії. Ми створили синтетичні дані, що відображають погодинне споживання енергії для 100 змодельованих домогосподарств, включаючи щоденні цикли використання та специфічні для домогосподарства закономірності. Проста нейронна мережа була навчена локально на даних кожного домогосподарства за допомогою стандартного методу оптимізації, а оновлення моделі були передані центральному серверу для покращення спільної глобальної моделі.

Результати та обговорення. Об'єднана модель досягла точності прогнозування, майже рівної традиційній централізованій моделі, зберігаючи при цьому конфіденційність даних. Ключовими факторами, що впливають на продуктивність, були частота навчання пристроїв локально перед обміном результатами та кількість домогосподарств, що брали участь у кожному раунді навчання. Підхід залишався точним, навіть коли лише половина пристроїв брала участь одночасно. Порівняно з неколаборативними моделями, які навчалися незалежно кожним домогосподарством, об'єднаний підхід запропонував суттєве покращення точності прогнозування. Ці результати показують, що хорошої продуктивності можна досягти, захищаючи конфіденційність користувачів та використовуючи прості моделі, придатні для пристроїв з низьким енергоспоживанням.

Висновки. Ця робота показує, що добре розроблене моделювання з реалістичними даними про споживання енергії може допомогти оцінити методи федеративного навчання за практичних обмежень. Навіть прості моделі, навчені децентралізованим способом із збереженням конфіденційності, можуть запропонувати корисні прогнози для інтелектуальних енергетичних систем. Цей підхід підходить для реального розгортання та може допомогти вдосконалити енергетичну аналітику з урахуванням конфіденційності.

Ключові слова: Федеративне навчання, інтелектуальні лічильники, прогнозування енергії, стохастичний градієнтний спуск, машинне навчання із збереженням конфіденційності, децентралізована оптимізація.

Received / Одержано	Revised / Доопрацьовано	Accepted / Прийнято	Published / Опубліковано
17 May, 2025	10 June, 2025	12 June, 2025	26 June, 2025

Electronics and information technologies, 2025, 30, 33–42
http://publications.lnu.edu.ua/collections/index.php/electronics/index

UDC 004.89

PARAMETER EFFICIENT FINE-TUNING AND OVERFITTING IN GPT LARGE LANGUAGE MODELS: A METRIC-BASED COMPARISON

Bohdan Pavlyshenko , ***Ivan Bulka*** *
Ivan Franko National University of Lviv,
50 Drahomanova St., 79005 Lviv, Ukraine

Pavlyshenko B. M, Bulka I. I. (2025). Parameter Efficient Fine-Tuning and Overfitting in GPT Large Language Models: a Metric-Based Comparison. *Electronics and Information Technologies*, 30, 33-42. https://doi.org/10.30970/eli.30.3

ABSTRACT

Background. Building upon previous research, this study conducts an exploration into Large Language Models (LLMs), with an emphasis on the fine-tuning and assessment of LLaMA-3.1 for instructional tasks. LLaMA-3.1, which is a new generation model and has gained considerable recognition based on its superior performance on various benchmarks. Besides assessing the disparities and improvements between the base and the fine-tuned versions of LLaMA-3.1 on an instruction dataset, the study also addresses the concern of overfitting with LLaMA-3.1. Furthermore, it carries out a comparison between LLaMA-3.1 and both its predecessor, LLaMA-2, and another LLM known as Mixtral, thereby providing a more comprehensive picture of LLaMA-3.1's capabilities compared to other models.

Materials and Methods. The fine-tuning of LLaMA-3.1 employed state-of-the-art techniques, such as Low-Rank Adaptation (LoRA) and Quantized Low-Rank Adaptation (QLoRA), on comprehensive instruction datasets. Acknowledging the resource-intensive nature of LLM fine-tuning, optimization measures were taken. The fine-tuning process was additionally enhanced using Parameter-Efficient Fine-tuning (PEFT) on NVIDIA A100 Tensor Core GPU (graphics processing unit) instances. All the models were fine-tuned using Hugging Face and PyTorch platforms for optimal performance.

Results and Discussion. The results obtained from fine-tuning and evaluating LLaMA-3.1 offer valuable insights into how this model performs with specific tasks. The evaluation framework proved helpful in the efficient assessment assessing LLMs' performance concerning instruction tasks. The research highlights the importance of evaluation for LLM applications. It shows that not always is fine-tuning a good choice, due to the nature of the model and the specifics of the task. It highlights the overfitting problem.

Conclusion. The close examination of LLaMA-3.1 contributes to the field of machine learning by offering insights into how this model works and its possible fine-tuning for special tasks. The findings of this research create opportunities for more in-depth studies around the application of LLMs. It highlights the importance of efficient evaluation with already designed metrics.

Keywords: LLMs, GPT, Mixtral, LLaMA, fine-tuning, overfitting

INTRODUCTION

The swift progress of Natural Language Processing (NLP) has been majorly steered by LLMs like Transformers [1, 2], Bidirectional Encoder Representations from Transformers (BERT) [3, 4], Generative Pretrained Transformer (GPT) [5], etc. They've cleared new ways for various tasks like text classification [6], machine translation [7], and summarization [8].

Now, more advanced models like GPT-3.5 [9], GPT-4 [10], and Claude [11] have expanded NLP's scope by simply following user instructions and explaining patterns.

LLMs can be helpful tools in media and communication, helping to distinguish between real news and fake or biased news [12]. They can also be used in finance, where they can be very helpful in conducting detailed studies of financial news [13, 14]. This widespread use of LLMs highlights their importance and the possibility of further study in different areas.

Applying LLMs to niche domains brings unique complications. Supervised Fine-Tuning [15] methods usually help tailor these LLMs for specific uses, but the balancing between providing comprehensive language capabilities and achieving sector-specific efficacy is complex. This difficulty becomes important in business settings where these models grapple with specialized queries needing custom solutions.

A new generation of models, like GPT-4, Claude, can be accessed via Application Programming Interfaces (APIs), which raises issues about private data handling. A large number of tasks can be solved using Retrieval Augmented Generation (RAG) [16]. However, the transfer of data to third-party apps is still inevitable. Increasing the number of API requests could also escalate costs.

Thus, the alternative would be to custom fine-tune [17] and store models on personal resources. This ensures data security and potentially offers cost advantages. LLMs can be used in various spaces like the media [18] and finance [13, 19], due to their broad range of applications.

This paper focuses on the Fine-tuning and evaluation of a new model created by Facebook. It's LLaMA-3.1 [20]. The model was fine-tuned using state-of-the-art methods like LoRA [21, 22] and QLoRA [23]. It compares the current model with fine-tuned models from the previous article [24].

We would focus on the latest LLaMA-3.1 and assess its performance in carrying out instructional tasks. This model has been thoroughly optimized by techniques like LoRA, QLoRA, and Parameter-Efficient Fine-Tuning (PEFT), which we validate through robust evaluation approaches. Our findings can guide future research and applications for Large Language Models.

MATERIALS AND METHODS

This study involves fine-tuning the model LLaMA-3.1 for instruction-based tasks. LLaMA-3.1 exists in three different sizes: 8B, 70B, and 405 B. Due to resource constraints and to compare results with the previous fine-tuned models, the 8b models were selected for this experiment. The PyTorch library was utilized for the fine-tuning process.

Training Dataset

A training dataset consolidates two freely accessible datasets, namely, Instruct-v3 [25] and Alpaca [26]. These datasets, crafted for refining instructions, are accessible from GitHub. To ensure suitability, a filtering process was conducted on these datasets to retain only those instructions composed of fewer than 1024 tokens.

A dataset was partitioned into three unique sections: training, validation, and testing. These sections contained 83k, 10k, and 3k records in their respective order. They served various purposes: the training section was used for refining the models, the validation section verified the efficacy of training during the refining process, and the testing section helped evaluate the efficiency of the final models. This dataset was used to fine-tune LLaMA-2 and Mixtral for the previous research. It allows us to compare LLaMA-3.1 with LLaMA-2 and Mixtral models that were fine-tuned in the previous paper.

A key element for fine-tuning LLaMA-3.1 is the formatting of the training dataset. Without proper formatting, the results may be significantly degraded and fail to reflect the true capabilities of the model. For this study, we employed a consistent template when preparing the dataset, ensuring that each sample followed the same conversational structure.

The dataset was serialized using a custom prompt template, designed to mimic a natural conversational exchange between a user and the assistant, while also allowing for the inclusion of system-level instructions. Each training sample in the dataset adheres to the following format:

```
<|begin_of_text|><|start_header_id|>system<|end_header_id|>
{{ system_prompt }}<|eot_id|><|start_header_id|>user<|end_header_id|>
{{ user_msg_1 }}<|eot_id|><|start_header_id|>assistant<|end_header_id|>
{{ model_answer_1 }}<|eot_id|>
```

In this notation, special tokens are used to denote the boundaries and roles within the conversation. The <|begin_of_text|> token marks the start of a new data sample. The <|start_header_id|> and <|end_header_id|> tokens enclose identifiers specifying the role of the content that follows (e.g., system, user, or assistant). The {{ system_prompt }}, {{ user_msg_1 }}, and {{ model_answer_1 }} placeholders are replaced with the actual system instruction, user input, and target assistant output, respectively. The <|eot_id|> token indicates the end of each segment or turn within the conversation. This structured formatting enables the model to clearly distinguish between instructions, queries, and responses, thereby enhancing learning effectiveness during fine-tuning.

LoRA and QLoRA settings

LoRA is a way to efficiently fine-tune LLMs by representing the Matrix of weights as a multiplication of 2 matrices with lower dimensions. The key element here is Matrix Rank (r), which affects the number of trainable parameters.

With the current matrix rank, 176 million parameters, which is 3.73% of all LLaMA-3.1 parameters. For the model, fine-tuning was used with LoraConfig (Table 1).

Table 1. **Lora Config for LLMs fine-tuning**

Parameter	Parameter description	Value
lora_alpha	LoRA scaling factor	16
lora_dropout	Dropout parameter to reduce overfitting	0.1
r	Matrix rank relates to the number of trainable parameters	64

For efficient comparison of LLaMA-3.1 with two previously trained models, those parameters were the same for all 3 models.

Training parameters

A model was tuned during 2 epochs. Considering that the base model captures lots of dependencies, a large number of epochs might cause overfitting [27]. Batch size is a parameter that represents the number of samples in the batch for training. We noticed that batch size can be increased for faster training, but compared to LLaMA-3.1 with previous experiments, we decided to use 4 as batch size [28]. The next parameters, as warmup_step (a way to reduce the primacy effect of the early training examples) [29], learning rate (indicate how fast a model could train) [30], 16-bit floating point format (represents QLoRA, quantization that helps to reduce the size of the model) [31].

LLaMA-3.1 could handle very big contexts (up to 128k tokens) [32], but a value of 1024 was selected to compare the current model with previous experiments. Training parameters can be found in Table 2.

Evaluation

The evaluation was done as in the previous experiment. To check how well the models worked, we used a test dataset that wasn't involved during the fine-tuning phase. First, we sent instructions, expected answers, and actual answers to the GPT-4 model, which then

Table 2. **Training parameters for LLMs fine-tuning**

Parameter	Parameter description	Value
num_train_epochs	Number of training epochs	2
per_device_train_batch_size	Batch size	4
warmup_steps	The number of warm-out steps	0.03
bf16	16-bit floating point format	True
max_seq_length	Max number of tokens	1024
learning_rate	Learning rate	2.5×10^{-5}

gave a score out of 10, with a higher score meaning better compliance with the instructions [24]. Second, we used the RAGAS [33] library to evaluate the models using two measures: Answer Correctness and Answer Semantic Similarity. You can access the RAGAS library via this link: https://docs.ragas.io/en/stable/.

RESULTS AND DISCUSSION

Training and validation loss comparison

Loss comparison is important to measure the efficiency of training or fine-tuning. For our experiments, we used the cross-entropy loss function [34], which is standard for tasks related to content generation. If loss decreases, that means that the model can train, capture patterns, and be more efficient in solving tasks related to the training dataset. We have noticed that during training, losses decreased from 1.09 to 1.06 for the LLaMA 3.1 model. It might be a good indicator. Train loss comparison can be found in Table 3.

Table 3. **Train loss comparison**

epoch	LLaMA-2	Mixtral	LLaMA-3.1
1	1.1736	1.0665	1.0887
2	1.1175	1.0723	1.0644

The most important is the validation loss. It's a value of the loss function on the dataset that was not used for training. It helps us to measure how good a model might be regarding data it had not seen previously. We noticed that validation losses also decreased (Table 4), which might be a good indicator too.

Table 4. **Validation loss comparison**

epoch	LLaMA-2	Mixtral	LLaMA-3.1
1	1.1474	1.0692	1.0756
2	1.1378	1.0626	1.0692

Training time comparison

A model was fine-tuned on an NVIDIA A100 Tensor Core GPU. Tuning LLMs requires a significant number of resources. Therefore, training time is also important as it impacts the cost of the solution.

We notice that LLaMA-3.1 requires 3 times more time for training than LLaMA-2 and 2.5 hours more than Mixtral (Table 5). Some techniques can significantly decrease training time, like Unsloth [35]. However, to ensure the fairness of the experiments and compatibility with previous research, it was decided to avoid current techniques.

Table 5. **Training time comparison**

Model	LLaMA-2	Mixtral	LLaMA-3.1
Training time, hours: mins	3:27	7:36	9:55

Metrics comparison

The key element for comparison between different Machine Learning algorithms is metric comparison. Evaluation was done on the testing dataset – a dataset that was not used for fine-tuning.

The evaluation of base LLaMA-3.1 against a fine-tuned model and models tuned in the previous paper demonstrates the importance of fine-tuning for specific tasks [24]. Three core metrics are used in this assessment: GPT-4 score Answer Correctness, and Answer Semantic Similarity [24]. The GPT-4 score is an automated evaluation metric in which the GPT-4 model is provided with the system instructions, user message, and both the golden (expected) and actual answers. Based on this information, GPT-4 assigns a score from 1 (worst) to 10 (best) that reflects how well the actual answer matches the golden one. Notably, each metric has its limitations, with Answer Semantic Similarity perhaps less suitable for specialized instruction tasks that may require knowledge from fields like physics or mathematics.

For comparison, the LLAMA-3.1 base outperforms other models, and the LLaMA-3.1 is fine-tuned (Table 6, Fig. 1).

That means that with fine-tuning, we overfit a model. Therefore, it can efficiently solve tasks related to training datasets and underperform on data it has not seen.

We noticed that LLaMA-3.1, base overperformed all previous models for all 3 metrics. Consequently, for the current task with the instructions dataset, the base model should be used.

Discussion and future direction

An empirical analysis of the LLaMA-2, LLaMA-3.1, and Mixtral models highlights their effectiveness in executing instructional tasks. Findings show that LLaMA-2 and LLaMA-3 tend to overfit. On the other hand, Mixtral outperforms LLaMA-2 during the evaluation phase, making it a more suitable option for instructional tasks. The key finding is that LLaMA-3.1 does not need fine-tuning to efficiently follow instructions. The base model works significantly better than the fine-tuned model.

We understand that for some domain-specific tasks, fine-tuning might be essential. General models were trained on a large amount of data, but can not know everything regarding specific domains. Also, often, those domains have some sensitive data, which makes it impossible to use models via API (like GPT-4 or Claude). Efficient fine-tuning is important for those domains.

Tuning models for RAG is also important, as the tuned model can answer user questions much better than the base model.

Table 6. **Metrics comparison**

Model	GPT-4 score (max 10)	Answer Correctness	Answer Semantic Similarity
LLaMA-2, base	7.21	0.66	0.91
LLaMA-2, fine-tuned	6.96	0.63	0.91
Mixtral, base	7.12	0.62	0.91
Mixtral, fine-tuned	7.51	0.67	0.91
LLaMA-3.1, base	8.54	0.72	0.92
LLaMA-3.1, fine-tuned	6.79	0.61	0.90

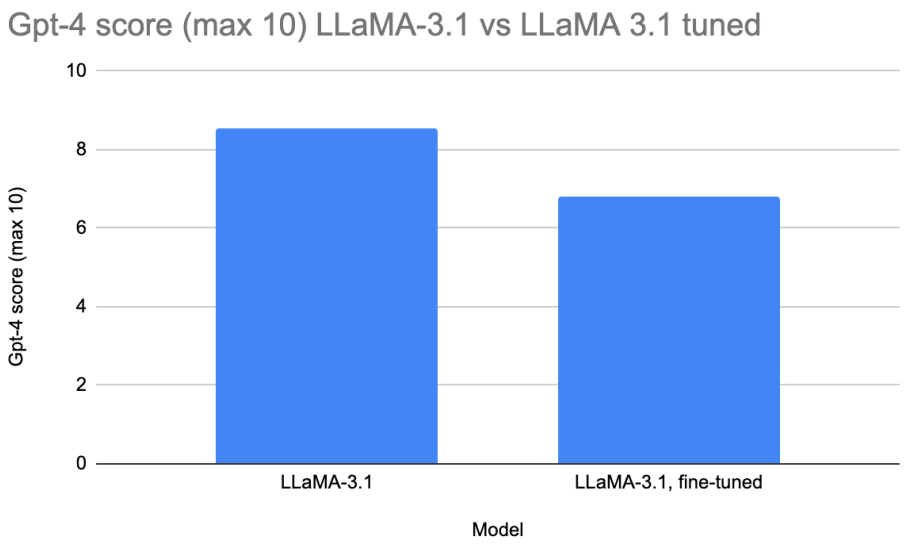


Fig. 1. Metrics comparison

We are going to test those approaches in different specific areas, not just teaching tasks or following instructions. We'll work to improve our methods and the way we measure performance in this study. This work might help us make better language models and push forward the field of language processing. It's also important to note that we're going to work a lot on getting better at the 'reasoning' part in future studies.

CONCLUSION

After an investigation of different Large Language Models (LLMs), particularly LLaMA-3.1, LLaMA-2, and Mixtral, our research has yielded interesting insights. Initially, Mixtral showed how great an impact fine-tuning can have on model performance. Similarly, we tried similar techniques on LLaMA-3.1, including cutting-edge methods like LoRA and QLoRA.

When fine-tuned with an instruction task, the performance of LLaMA-3.1 took a hit. LLaMA-3.1, in its basic form, is already very good at solving instruction tasks, and the addition of fine-tuning brought on too much specialization, leading to overfitting. Despite this, we believe that for some particular tasks, especially within specific domains, fine-tuning might still be necessary to achieve enhanced performance.

Upon comparing LLaMA-2 and Mixtrail, we observed that LLaMA-2 fine-tuned faster but was more susceptible to overfitting. Mixtrail, although slower in training, proved to be consistently better at handling instructional tasks in our tests, suggesting it has a better balance between general language skills and specificity.

This tendency towards overfitting, observed in the LLaMA family, helps us understand the importance of carefully managing the tuning process. We recommend cutting down on tuning epochs to prevent overfitting.

Our research leveraged the RAGAS library to evaluate LLM performance, a practice we think would be instrumental in future machine learning studies where LLMs are used. Our conclusions provide crucial learning about the workings of LLaMA-3.1 and LLMs in general, including their performance, fine-tuning practices, and predisposition towards overfitting.

Acquired knowledge opens the gate for future studies on LLMs, which we believe is potentially pivotal for unlocking their full capability, particularly concerning specific tasks and domains. We're also reminded of the critical part fine-tuning plays in amplifying LLM

performance and the necessity of finding a good balance between general language competencies and specific task efficiency, especially in domain-specific tasks.

AUTHOR CONTRIBUTIONS

Conceptualization, [*B.P., I.B.*]; methodology, [*I.B.*]; validation, [*B.P.*]; formal analysis, [*I.B.*.]; investigation, [*I.B.*]; resources, [*I.B.*]; data curation, [*I.B.*]; writing – original draft preparation, [*I.B.*]; writing – review and editing, [*B.P.*]; visualization, [*I.B.*] supervision, [*B.P.*]; project administration, [*I.B.*].

All authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30.

[2] De Leon, Francisco, Pablo Gómez, Juan A. Martinez-Velasco, and Michel Rioual. "Transformers." In Power system transients, pp. 177-250. CRC Press, 2017

[3] Rogers, A., Kovaleva, O., & Rumshisky, A. (2021). A primer in BERTology: What we know about how BERT works. *Transactions of the association for computational linguistics*, *8*, 842-866. https://doi.org/10.1162/tacl_a_00349

[4] Hao, Y., Dong, L., Wei, F., & Xu, K. (2019). Visualizing and understanding the effectiveness of BERT. *arXiv preprint arXiv:1908.05620*. https://doi.org/10.48550/arXiv.1908.05620

[5] Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., ... & McGrew, B. (2023). Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*. https://doi.org/10.48550/arXiv.2303.08774

[6 Liga, D., & Robaldo, L. (2023). Fine-tuning GPT-3 for legal rule classification. *Computer Law & Security Review*, *51*, 105864. https://doi.org/10.1016/j.clsr.2023.105864

[7] Hendy, A., Abdelrehim, M., Sharaf, A., Raunak, V., Gabr, M., Matsushita, H., ... & Awadalla, H. H. (2023). How good are gpt models at machine translation? a comprehensive evaluation. *arXiv preprint arXiv:2302.09210*. https://doi.org/10.48550/arXiv.2302.09210

[8] Goyal, T., Li, J. J., & Durrett, G. (2022). News summarization and evaluation in the era of gpt-3. *arXiv preprint arXiv:2209.12356*. https://doi.org/10.48550/arXiv.2209.12356

[9] Koubaa, A. (2023). GPT-4 vs. GPT-3.5: A concise showdown. *Preprints. org*, *2023030422*.

[10] Mao, R., Chen, G., Zhang, X., Guerin, F., & Cambria, E. (2023). GPTEval: A survey on assessments of ChatGPT and GPT-4. *arXiv preprint arXiv:2308.12488*. https://doi.org/10.48550/arXiv.2308.12488

[11] Adetayo, A. J., Aborisade, M. O., & Sanni, B. A. (2024). Microsoft Copilot and Anthropic Claude AI in education and library service. *Library Hi Tech News*. https://doi.org/10.1108/LHTN-01-2024-0002

[12] Pavlyshenko, B. M. (2023). Analysis of disinformation and fake news detection using fine-tuned large language model. *arXiv preprint arXiv:2309.04704*. https://doi.org/10.48550/arXiv.2309.04704

[13] Pavlyshenko, B. M. (2023). Financial news analytics using fine-tuned llama 2 gpt model. *arXiv preprint arXiv:2308.13032*. https://doi.org/10.48550/arXiv.2308.13032

[14] Pavlyshenko, B. M. (2022). Methods of informational trends analytics and fake news detection on twitter. *arXiv preprint arXiv:2204.04891*. https://doi.org/10.48550/arXiv.2204.04891

[15] Prottasha, N. J., Sami, A. A., Kowsher, M., Murad, S. A., Bairagi, A. K., Masud, M., & Baz, M. (2022). Transfer learning for sentiment analysis using BERT based supervised fine-tuning. *Sensors*, *22*(11), 4157. https://doi.org/10.3390/s22114157
[16] Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., ... & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in neural information processing systems*, *33*, 9459-9474.
[17] Lin, X., Wang, W., Li, Y., Yang, S., Feng, F., Wei, Y., & Chua, T. S. (2024, July). Data-efficient Fine-tuning for LLM-based Recommendation. In *Proceedings of the 47th international ACM SIGIR conference on research and development in information Retrieval* (pp. 365-374). https://doi.org/10.1145/3626772.365780
[18] Meier, R. (2024). Llm-aided social media influence operations. *Large Language Models in Cybersecurity: Threats, Exposure and Mitigation*, 105-112. https://doi.org/10.1007/978-3-031-54827-7_11
[19] Yu, X., Chen, Z., Ling, Y., Dong, S., Liu, Z., & Lu, Y. (2023). Temporal data meets LLM--explainable financial time series forecasting. *arXiv preprint arXiv:2306.11025*. https://doi.org/10.48550/arXiv.2306.11025
[20] Vavekanand, R., & Sam, K. (2024). Llama 3.1: An in-depth analysis of the next-generation large language model. https://doi.org/10.48550/arXiv.2306.11025
[21] Li, Y., Yu, Y., Liang, C., He, P., Karampatziakis, N., Chen, W., & Zhao, T. (2023). Loftq: Lora-fine-tuning-aware quantization for large language models. *arXiv preprint arXiv:2310.08659*. https://doi.org/10.48550/arXiv.2310.08659
[22] Wu, B., Zhu, R., Zhang, Z., Sun, P., Liu, X., & Jin, X. (2024). {dLoRA}: Dynamically orchestrating requests and adapters for {LoRA}{LLM} serving. In *18th USENIX Symposium on Operating Systems Design and Implementation (OSDI 24)* (pp. 911-927).
[23] Zhang, X., Rajabi, N., Duh, K., & Koehn, P. (2023, December). Machine translation with large language models: Prompting, few-shot learning, and fine-tuning with QLoRA. In Proceedings of the Eighth Conference on Machine Translation (pp. 468-481).
[24] Pavlyshenko, B., & Bulka, I. (2024) Metric-Based Comparison of Fine-Tuned LLAMA 2 and MIXTRAL Large Language Models for Instruction Tasks. *Electronics and information technologies/Електроніка та інформаційні технології*, 26, 16-24. http://dx.doi.org/10.30970/eli.26.2
[25] MosaicML. (2023). MosaicML Instruct-v3 Dataset. Hugging Face. https://huggingface.co/datasets/mosaicml/instruct-v3
[26] Rohan Taori and Ishaan Gulrajani and Tianyi Zhang and Yann Dubois and Xuechen Li and Carlos Guestrin and Percy Liang and Tatsunori B. Hashimoto. (2023). Stanford Alpaca: An Instruction-following LLaMA model. Hugging Face. https://huggingface.co/datasets/tatsu-lab/alpaca
[27] Ying, X. (2019, February). An overview of overfitting and its solutions. In *Journal of physics: Conference series* (Vol. 1168, p. 022022). IOP Publishing. https://doi.org/10.1088/1742-6596/1168/2/022022
[28] You, Y., Wang, Y., Zhang, H., Zhang, Z., Demmel, J., & Hsieh, C. J. (2020). The limit of the batch size. *arXiv preprint arXiv:2006.08517*. https://doi.org/10.48550/arXiv.2006.08517
[29] Kang, F., Just, H. A., Sun, Y., Jahagirdar, H., Zhang, Y., Du, R., ... & Jia, R. (2024). Get more for less: Principled data selection for warming up fine-tuning in llms. *arXiv preprint arXiv:2405.02774*. https://doi.org/10.48550/arXiv.2405.02774
[30] Sanyal, S., Neerkaje, A., Kaddour, J., Kumar, A., & Sanghavi, S. (2023). Early weight averaging meets high learning rates for llm pre-training. *arXiv preprint arXiv:2306.03241*. https://doi.org/10.48550/arXiv.2306.03241
[31] Agrawal, A., Mueller, S. M., Fleischer, B. M., Sun, X., Wang, N., Choi, J., & Gopalakrishnan, K. (2019, June). DLFloat: A 16-b floating point format designed for

deep learning training and inference. In *2019 IEEE 26th Symposium on Computer Arithmetic (ARITH)* (pp. 92-95). IEEE. https://doi.org/10.1109/ARITH.2019.00023

[32] Grattafiori, A., Dubey, A., Jauhri, A., Pandey, A., Kadian, A., Al-Dahle, A., ... & Vasic, P. (2024). The llama 3 herd of models. *arXiv preprint arXiv:2407.21783*. https://doi.org/10.48550/arXiv.2407.21783

[33] Es, S., James, J., Anke, L. E., & Schockaert, S. (2024, March). Ragas: Automated evaluation of retrieval augmented generation. In *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations* (pp. 150-158).

[34] VM, K., Warrier, H., & Gupta, Y. (2024). Fine tuning llm for enterprise: Practical guidelines and recommendations. *arXiv preprint arXiv:2404.10779*. https://doi.org/10.48550/arXiv.2404.10779

[35] Zheng, Y., Zhang, R., Zhang, J., Ye, Y., Luo, Z., Feng, Z., & Ma, Y. (2024). Llamafactory: Unified efficient fine-tuning of 100+ language models. *arXiv preprint arXiv:2403.13372*. https://doi.org/10.48550/arXiv.2403.13372

ЕФЕКТИВНЕ ТОЧНЕ НАЛАШТУВАННЯ ПАРАМЕТРІВ ТА ПЕРЕНАВЧАННЯ В МОВНИХ МОДЕЛЯХ GPT: ПОРІВНЯННЯ НА ОСНОВІ МЕТРИК

Богдан Павлишенко **, *Іван Булка*** ***
Львівський національний університет імені Івана Франка,
вул. Драгоманова 50, 79005 Львів, Україна

АНОТАЦІЯ

Вступ. Спираючись на попередні дослідження, це дослідження зосереджується на великих мовних моделях (LLM) з фокусом на тонкому налаштуванні та оцінці LLaMA-3.1 для завдань зв'язаних з інструкціями. LLaMA-3.1, яка є моделлю нового покоління і здобула значне визнання завдяки своїм чудовим результатам. Окрім оцінки відмінностей і вдосконалень між базовою та налаштованою версіями LLaMA-3.1 на наборі даних інструкцій, дослідження також звертає увагу на проблему перенавчання LLaMA-3.1. Додатково було проведено порівнянням між LLaMA-3.1, її попереднцею, LLaMA-2, а також іншою LLM, відомою як Mixtral, що дозволяє отримати більш повну картину можливостей LLaMA-3.1.

Матеріали та методи. Для тонкого налаштування LLaMA-3.1 використовувались сучасні підходи, такі як адаптація низького рангу (LoRA) і квантована адаптація низького рангу (QLoRA), на комплексних наборах даних інструкцій. Враховуючи ресурсоємність процесу тонкого налаштування LLM, вживались заходи щодо його оптимізації. Процес тонкого налаштування був удосконалений за допомогою Параметрично ефективного тонкого налаштування (PEFT) на екземплярах NVIDIA A100 Tensor Core GPU. Усі моделі були налаштовані за допомогою платформ Hugging Face і PyTorch для досягнення найкращої продуктивності. Дослідження підкреслює важливість ретельної оцінки LLM для практичних застосувань.

Результати. Результати, отримані в результаті тонкого налаштування та оцінки LLaMA-3.1, надали цінну інформацію про те, як ця модель виконує конкретні завдання. Система оцінювання виявилася корисною для ефективної оцінки ефективності LLM на завданнях з інструкціями. Показано, що точне налаштування не завжди є найкращим

вибором з огляду на специфіку моделі та особливості завдання. Дослідження підкреслює проблему перенавчання в LLM.

Висновки. Ретельний аналіз LLaMA-3.1 робить внесок у сферу машинного навчання, поглиблюючи розуміння особливостей роботи цієї моделі та можливостей її тонкого налаштування для конкретних завдань. Результати цього дослідження створюють підгрунтя для подальших досліджень і застосування LLMs та підкреслюють значення ефективної оцінки з використанням існуючих метрик.

Ключові слова: LLMs, GPT, Mixtral, LLaMA, тонке налаштування, перенавчання.

Received / Одержано	Revision / Доопрацьовано	Accepted / Прийнято	Published / Опубліковано
18 April, 2025	02 May, 2025	08 May, 2025	26 June, 2025

Electronics and Information Technologies, 2025, 30, 43–58
http://publications.lnu.edu.ua/collections/index.php/electronics/index

UDC: 004.89 + 004.738.5

SEMANTIC SIMILARITY ANALYSIS USING TRANSFORMER-BASED SENTENCE EMBEDDINGS

Bohdan Pavlyshenko **, *Mykola Stasiuk*** ***
Ivan Franko National University of Lviv,
50 Drahomanova St., 79005 Lviv, Ukraine

Pavlyshenko B. M., Stasiuk M. I. (2025). Semantic Similarity Analysis Using Transformer-Based Sentence Embeddings. *Electronics and Information Technologies*, 30, 43–58.
https://doi.org/10.30970/eli.30.4

ABSTRACT

Background. Transformer-based models have become central to natural language processing, demonstrating state-of-the-art performance in semantic similarity assessment, a task critical for various applications. These models capture detailed relationships between text, advancing the ability to gauge semantic relatedness.

Materials and Methods. The performance of sentence embedding models, including *all-mpnet-base-v2*, *all-MiniLM-L6-v2*, *paraphrase-multilingual-mpnet-base-v2*, *bge-base-en-v1.5*, *all-roberta-large-v1*, *all-distilroberta-v1*, *LaBSE*, *paraphrase-MiniLM-L3-v2*, *bge-large-en-v1.5*, was assessed across different dataset sizes with two datasets. The following preprocessing steps were applied to the datasets: lowercasing, removing stop words, cleaning from special symbols and numbers, and lemmatization. Cosine similarity scores with negative values, indicating semantic dissimilarity, were treated as equivalent to a human-annotated similarity score of 0, and non-negative cosine similarity values were scaled to the 0-5 range. Metrics such as R^2, MSE, RMSE, MAE, Spearman's Correlation Coefficient, and Kendall's Tau were used for evaluation.

Results and Discussion. Models' performance generally improves with increased data. Evaluation of sentence embedding models revealed performance variations. *all-roberta-large-v1* showed strong accuracy with high R^2values and low errors. *BAAI/bge-large-en-v1.5* excelled in capturing semantic relationships, demonstrating high Spearman's and Kendall's Tau coefficients. *all-MiniLM-L6-v2* demonstrated the fastest embedding generation. *BAAI/bge-base-en-v1.5* presented the lowest accuracy. Processing times generally increase with data size.

Conclusion. This study highlights a trade-off between accuracy and efficiency in sentence embedding. Model selection depends on balancing these factors to align with specific application needs. In cases when requiring high accuracy should favor *all-roberta-large-v1*, while those prioritizing speed would benefit from *all-MiniLM-L6-v2*. *BAAI/bge-large-en-v1.5* is most suitable for tasks demanding semantic understanding of text details.

Keywords: semantic similarity, sentence embeddings, transformers.

INTRODUCTION

As digital information grows at an accelerating pace, the ability to understand the meaning behind text has become fundamental. From search engines to nuanced interactions with virtual assistants, the need for machines to comprehend semantic content is becoming increasingly important. However, while computers excel at processing raw textual data, catching the hidden nuances of natural language remains a complex

challenge. The core of this challenge lies in making it possible for machines to move beyond simple word matching and to understand the underlying meaning of text instead.

Researchers have developed text embeddings [1], a powerful technique that transforms words and sentences into dense numerical vectors to address this challenge. Unlike simple word representations, these vectors capture the semantic relationships between text elements, allowing computers to understand context and meaning. This process has evolved significantly, moving from earlier methods like Word2Vec [2] and GloVe [3], which generated static word embeddings, to more advanced contextual embeddings such as those produced by transformers [4] and transformer-based models like BERT [5] and its variants. These advancements have paved the way for the development of various application areas for text embeddings, ranging from fundamental tasks like text classification [6] to more sophisticated techniques such as data augmentation using large language models [7]. The benefit of converting text to vectors is that mathematical operations such as calculating distance can be performed on them, thus enabling computers to quantify semantic relationships [8, 9].

A crucial application of text embeddings is the semantic similarity measurement, quantifying the degree to which two pieces of text share a related meaning. This capability enables a wide range of Natural Language Processing tasks. For instance, accurate semantic similarity in information retrieval allows search engines to identify documents conceptually relevant to a user's query, even if they do not contain the exact words [10]. Similarly, question-answering systems enable matching questions to appropriate answers by understanding their underlying meaning [11]. Other applications include plagiarism detection, where subtle similarities in phrasing can be identified, and document clustering, where related documents are grouped based on their semantic content [12]. The performance of all these applications is directly tied to the quality and accuracy of the text embeddings used to calculate semantic similarity.

Building on these advancements, the broader landscape of Natural Language Processing has recently been transformed by powerful Large Language Models (LLMs) such as Llama 3/4, Grok, Qwen, or Mistral. While these general-purpose LLMs demonstrate impressive capabilities in broad generative and complex reasoning tasks, it is crucial to differentiate them from specialized sentence embedding models. The models primarily evaluated in this study are specifically engineered and optimized for generating high-quality, dense vector representations of text, making them inherently more efficient and effective for the semantic similarity tasks under investigation.

While numerous sentences embedding models have been introduced recently, direct comparisons of their performance and efficiency are not always prevalent. This article therefore, undertook a comparative investigation of various sentence embedding models to address the challenge of accurately quantifying semantic similarity between texts. Specifically, the output of these models, when paired with cosine similarity, was analyzed to estimate the degree of semantic relatedness between different text snippets. Furthermore, to provide a practical perspective on the efficiency of these models, the time taken by each model to generate embeddings and subsequently compute the similarity scores was also measured and compared. This dual approach offered insights into the effectiveness and computational cost of employing different sentence embedding techniques for text similarity tasks.

Obtained insights are not merely theoretical; they directly inform the practical selection of models for diverse real-world applications. For instance, in scenarios demanding high-throughput processing with limited computational resources, such as real-time conversational AI, where models provide rapid, contextually-aware responses by semantically matching user input to relevant intents, or large-scale document indexing, where documents are transformed into searchable embeddings to enable finding content by meaning and efficient content organization, understanding a model's efficiency becomes paramount. Conversely, for applications where strong predictive capability and nuanced semantic understanding are critical, such as precise information retrieval for identifying the

most relevant passages to a user's query or complex question-answering systems that must accurately extract answers from vast knowledge bases, model effectiveness takes precedence. The quality of these generated embeddings is therefore paramount for the success of such systems. This study systematically quantifies these crucial trade-offs, providing an empirical basis for developers and researchers to choose the most suitable sentence embedding model among those utilized, taking into account considerations such as efficiency and predictive performance.

MATERIALS AND METHODS

In experiments, two datasets, publicly available on the Hugging Face Portal and designed explicitly for semantic similarity tasks, were utilized: sts-companion [13] and mteb-stsbenchmark-sts [14]. These datasets contain pairs of sentences along with a corresponding human-annotated semantic similarity score. Both datasets use similarity scores, which lie on a scale of 0 to 5 and represent varying degrees of semantic relatedness.

One of the datasets consists of 5289 records, and the other 8628. Both datasets were tested with initial subsets of 100, 500, 1000, 2000, 4000, and 5000 records. The larger dataset was also tested with 8000, and the full dataset size, while the smaller dataset was tested with the full dataset size. This approach allowed us to assess how model effectiveness varies with an increasing number of records.

To evaluate the effectiveness of different sentence embedding models for text similarity comparison, we employed the following models, all of which are available on the Hugging Face portal:

- sentence-transformers/all-MiniLM-L6-v2: this model is a compact and efficient Transformer-based architecture. It is designed to produce effective sentence embeddings with relatively small model size, making it suitable for applications where computational resources are limited or speed is crucial;
- *sentence-transformers/paraphrase-multilingual-mpnet-base-v2*: this model leverages the MPNet architecture [15] and has been trained on a large multilingual corpus. It is fine-tuned explicitly for paraphrase identification and semantic similarity across various languages, making it a strong choice for cross-lingual text understanding tasks;
- *BAAI/bge-base-en-v1.5*: BAAI developed this model based on the transformer architecture and trained it on a substantial English language dataset. It is recognized for achieving high performance on various text embedding benchmarks, demonstrating its ability to capture nuanced semantic relationships in English text;
- *sentence-transformers/all-roberta-large-v1*: this model utilizes the RoBERTa architecture [16], a powerful transformer variant known for its strong language understanding capabilities. Being a "large" model, it has a greater capacity to learn complex semantic representations, often leading to high accuracy in sentence embedding tasks;
- *sentence-transformers/all-distilroberta-v1*: this model is a distilled version of RoBERTa, meaning it retains much of the performance of its larger counterpart while having a smaller size and faster inference speed. It offers a good balance between accuracy and efficiency for generating sentence embeddings;
- *sentence-transformers/LaBSE*: standing for Language-Agnostic BERT Sentence Embeddings [17], LaBSE is explicitly designed to produce embeddings comparable across many different languages. This makes it highly effective for multilingual semantic similarity tasks and cross-lingual information retrieval;
- *sentence-transformers/paraphrase-MiniLM-L3-v2*: like the *all-MiniLM-L6-v2* model, this is another efficient transformer-based model focused on generating sentence embeddings. The "paraphrase" in its name indicates that it has been fine-tuned explicitly on paraphrase datasets, making it well-suited for tasks involving semantic equivalence;

- *sentence-transformers/all-mpnet-base-v2:* is based on the MPNet architecture, which is known for its effectiveness in capturing semantic relationships between sentences. It is designed to produce high-quality sentence embeddings and demonstrates a good balance of accuracy and computational efficiency. This model is suitable for a wide range of semantic similarity tasks.
- *BAAI/bge-large-en-v1.5*: is the larger variant of BAAI's *bge-base-en-v1.5* model for English. With more parameters, it aims to present state-of-the-art performance on English language text embedding tasks. It offers higher accuracy than its base counterpart at the cost of increased computational resources.

For this study, all models were evaluated exclusively on English language texts, irrespective of their inherent multilingual capabilities.

Fig. 1 illustrates the characteristics of the datasets used in this study. The top row displays the distribution of similarity scores for the mteb/stsbenchmark-sts (left) and sts-companion (right) datasets. The x-axis represents the similarity score, while the y-axis represents the frequency of occurrence. The distributions reveal differences in the two datasets' range and concentration of similarity scores. While both datasets have many texts with similar lengths, their score distributions differ. The mteb/stsbenchmark-sts dataset has a more uniform distribution of scores, whereas the sts-companion dataset shows a higher concentration of high-value scores.

Figure 1's middle and bottom rows show the distributions of sentence lengths for the first and second sentences in the comparison pairs, respectively. The x-axis represents the sentence length (number of tokens), and the y-axis represents the frequency of occurrence. These plots provide information about the variability in sentence lengths within each dataset. Notably, sts-companion contains a broader range of sentence lengths than mteb/stsbenchmark-sts, potentially impacting sentence embedding models' computational demands and performance.

To quantitatively evaluate the performance of each sentence embedding model in the semantic similarity task, we employed four commonly used regression metrics:

- **Mean Squared Error (MSE):** calculates the average of the squared differences between the predicted and the actual similarity scores. Lower MSE values indicate better model performance.
- **Root Mean Squared Error (RMSE)**: the RMSE is the square root of the MSE. It measures the average magnitude of the errors in the same units as the target variable, making it more interpretable than MSE.
- **Mean Absolute Error (MAE)**: calculates the average absolute differences between the predicted and actual similarity scores. Like RMSE, MAE is expressed in the units of the target variable and provides a robust measure of error, less sensitive to outliers than MSE or RMSE.
- $\boldsymbol{R^2}$**:** the coefficient of determination measures the proportion of the variance in the human-annotated similarity scores explained by the predicted scores. R^2 values range from 0 to 1, with higher values indicating a better fit of the model to the data.
- **Spearman's Rank Correlation Coefficient**: assesses the strength and direction of the monotonic association between the predicted and actual similarity values. It quantifies how well the ranking of predicted ratings aligns with human-annotated scores. Spearman's correlation spans from –1 to 1, where figures approaching 1 denote a strong positive monotonic link, figures nearing –1 suggest a strong negative monotonic link, and values close to 0 imply a weak or nonexistent monotonic link.
- **Kendall's Tau**: measures the ordinal association between the predicted and actual similarity scores. It evaluates the degree to which the predicted ordering agrees with the human-annotated ordering. Kendall's Tau figures range from –1 to 1, with results approaching 1 indicating strong agreement in the order, results nearing –1 indicating strong disagreement, and results near zero indicating weak or no agreement.

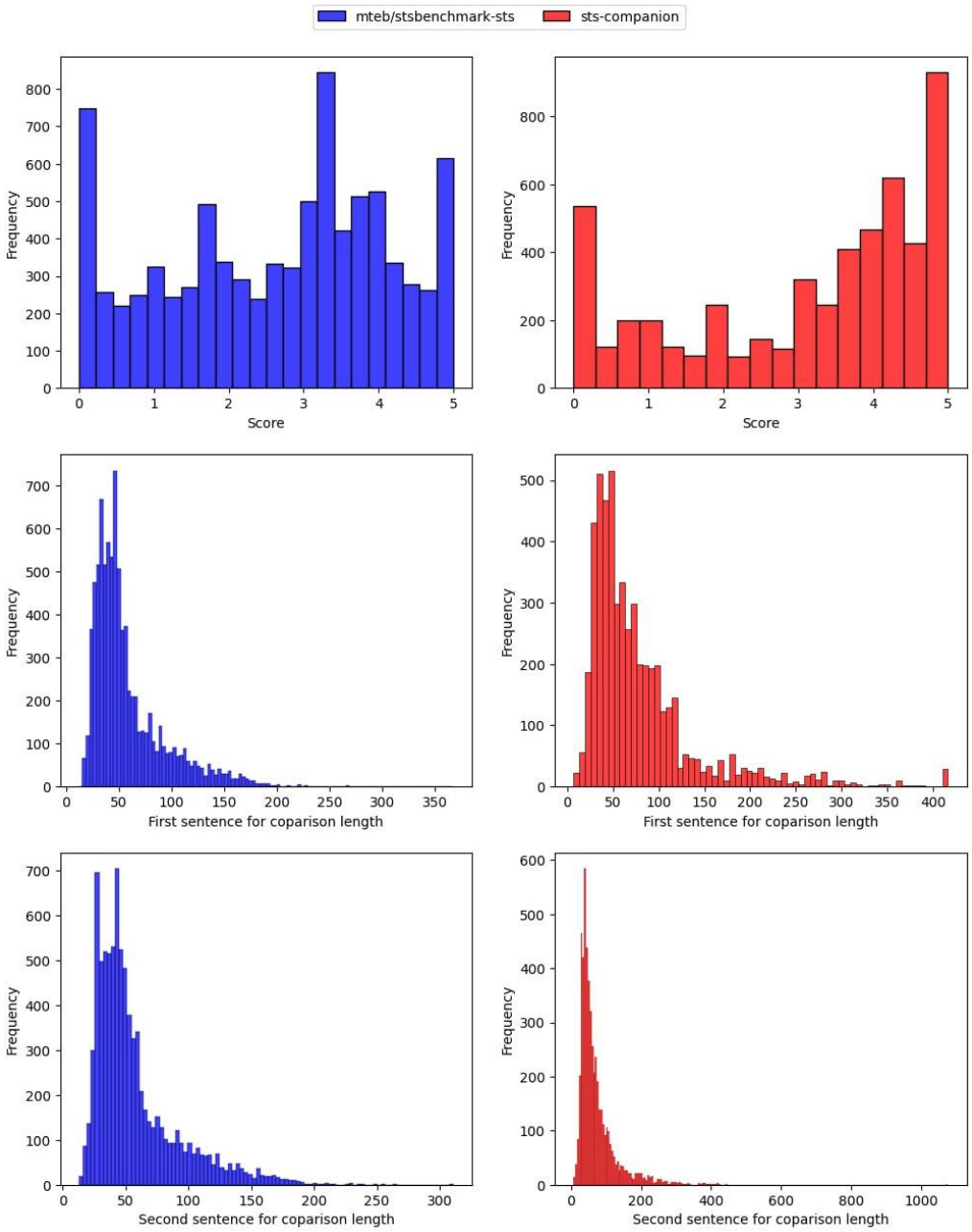


Fig. 1. Distribution of similarity scores and sentence lengths in the mteb/stsbenchmark-sts and sts-companion datasets.

These metrics collectively provide a comprehensive assessment of the sentence embedding models' performance in predicting semantic similarity. Additionally, we measured the time each model took to embed the text and calculated the average of these times to compare the computational efficiency of the models.

The experimental procedure involved the following steps.

1. The sentence pairs from each dataset were preprocessed to ensure consistency. The preprocessing included: lowercasing, removing stop words, cleaning from special symbols and numbers, and lemmatization;

2. Each sentence in the paired data was encoded into a vector embedding using the selected sentence embedding models.
3. The semantic similarity between the resulting vector embeddings for each sentence pair was calculated using cosine similarity. It is important to note that cosine similarity yields a score between –1 and 1. For this study, negative cosine similarity values, indicating semantic dissimilarity, were treated as equivalent to a human-annotated similarity score of 0, and only non-negative cosine similarity values were scaled to the 0-5 range.
4. The scaled cosine similarity scores were compared against the ground truth similarity scores to assess the effectiveness of each sentence embedding model in capturing semantic relationships.

RESULTS AND DISCUSSION

In this study, we made a simplifying assumption regarding the interpretation of negative cosine similarity values. Specifically, any sentence pair with negative cosine similarity calculated from the model-generated embeddings was assigned a scaled similarity score corresponding to the lowest possible human-annotated score - 0. This decision was made to align the model output with the lower bound of the human scoring scale and to focus the analysis primarily on the model's ability to capture positive semantic similarity. This approach is often adopted as a pragmatic choice when mapping model output to human-annotated scales that typically do not assign granular scores below semantic unrelatedness, i.e., below zero similarity. It simplifies the evaluation framework by focusing on the strength of positive semantic relationships, which is often the primary objective in tasks such as information retrieval and question answering.

However, it is important to acknowledge that this approach has limitations. We lose some of the information provided by the model's embedding space by treating all negative cosine similarity values as equivalent. A highly negative cosine similarity suggests a more substantial degree of semantic dissimilarity, potentially even indicating some semantic opposition, than a slightly negative value. Our simplification collapses these distinctions. Furthermore, this approach might mask subtle differences in how the models represent semantic dissimilarity. Some models might consistently produce more negative values for unrelated or contrasting sentences, and this behavior is not captured in our scaled scores. Future work could explore alternative scaling methods, such as mapping the full range of cosine similarity to the 0 - 5 scale, to better preserve the information contained in negative similarity values and provide a more detailed comparison of model performance.

Table 1 presents the times taken by each model to generate text embeddings for text from the mteb-stsbenchmark-sts dataset. The *paraphrase-MiniLM-L3-v2* model demonstrates the most efficient performance, with embedding times around 2 ms for smaller subsets, reaching only 3.7 ms for the whole dataset. The next best performance was exhibited by *all-MiniLM-L6-v2* and *all-distilroberta-v1* models; their times are approximately 3 ms for smaller subsets. However, *all-distilroberta-v1* climbs to 18.6 ms for the complete dataset, while *all-MiniLM-L6-v2* attains 5.6 ms for the most extensive dataset. Similar performance for smaller subsets was shown by *bge-base-en-v1.5* and *paraphrase-multilingual-mpnet-base-v2* – around 4–5 ms; for the 8628-record dataset, times grow to 34.8 ms and 38.8 ms, respectively. The *all-mpnet-base-v2* model has times of roughly 5–7 ms for small subsets and 35.1 ms for the largest. The *LaBSE* model's time expands to 36.3 ms for the whole dataset, compared to about 5–6 ms for smaller ones. In contrast, *all-roberta-large-v1* and *bge-large-en-v1.5* are the slowest; for the 8628-record dataset, times surge to 142.7 ms and 162.6 ms, respectively. While most models maintain relatively stable embedding times for smaller subsets (100 to 2000 records*), all-roberta-large-v1* and *bge-large-en-v1.5* display a substantial jump for the whole dataset.

Table 1. **Times taken by every model to create a vector representation of strings from the *mteb-stsbenchmark-sts* dataset.**

Model \ Records number	Time, ms							
	100	500	1000	2000	4000	5000	8000	8628
all-mpnet-base-v2	7.3	5.3	5.3	5.3	5.3	5.2	33.9	35.1
all-MiniLM-L6-v2	3.1	3.0	3.0	3.0	3.1	3.1	4.5	5.6
paraphrase-multilingual-mpnet-base-v2	4.7	4.9	4.9	4.8	5.0	4.7	38.5	38.8
bge-base-en-v1.5	4.7	4.7	4.7	4.6	4.9	4.6	33.4	34.8
all-roberta-large-v1	9.7	8.4	8.2	8.3	56.2	8.4	141.4	142.7
all-distilroberta-v1	3.3	3.0	3.1	3.0	17.5	3.0	18.3	18.6
LaBSE	6.2	4.8	4.7	4.8	32.3	4.6	35.6	36.3
paraphrase-MiniLM-L3-v2	2.9	2.0	2.1	2.2	3.2	2.2	3.0	3.7
bge-large-en-v1.5	10.1	8.3	8.3	8.4	123.0	8.3	132.8	162.6

For most models, the time to embed a given number of records remains relatively stable across the smaller subset sizes (100 to 2000 records). This suggests that the embedding process scales consistently for these models within this range. However, the time taken for some models noticeably expands when processing the larger subsets, particularly the full dataset. For example, the embedding time for *all-roberta-large-v1* surges to 142.7 ms for the 8628-record dataset, compared to approximately 8.3 ms for smaller subsets. Similarly, *bge-large-en-v1.5* also exhibits a substantial rise, reaching 162.6 ms for the most extensive dataset.

In contrast, models like *all-MiniLM-L6-v2* and *paraphrase-MiniLM-L3-v2* show the smallest increase in processing time when moving to the full dataset, suggesting more consistent scaling performance. *LaBSE* also jumps to 36.3 ms for the most extensive dataset. This suggests that the computational cost of processing the full dataset introduces additional overhead for some models but not all, highlighting differences in how efficiently these models handle larger inputs.

Table 2 shows the performance of each model in terms of embedding time. The *paraphrase-MiniLM-L3-v2* model demonstrates the most efficient performance, consistently achieving the fastest embedding times across all dataset sizes, with times around 2 ms. In contrast, the *bge-large-en-v1.5* model is the slowest, with embedding times growing with the dataset size and reaching 187.2 ms for the largest dataset of 5289 records. This significant rise suggests that its computational demands scale more substantially with larger inputs. Several other models also show relatively fast performance across the smaller dataset sizes, with times generally below 6 ms, including *all-mpnet-base-v2*, *paraphrase-multilingual-mpnet-base-v2*, and *bge-base-en-v1.5*. However, their embedding times display a noticeable climb for the largest dataset. Specifically, *all-roberta-large-v1* and *LaBSE* attain a moderate expansion, reaching 50.1 ms and 45.7 ms, respectively, for the 5289-record dataset. *all-distilroberta-v1* also shows a jump at the largest dataset size, reaching 22.6 ms.

Most models exhibit relatively consistent embedding times across varying dataset sizes. This suggests that the processing demands remain stable for these models. However, some models experience increased processing duration when handling the largest dataset of 5289 records. For instance, the embedding time for *all-roberta-large-v1*

Table 2. **Times taken by every model to create a vector representation of strings from the *sts-companion* dataset.**

	Time, ms						
Records number / Model	100	500	1000	2000	4000	5000	5289
all-mpnet-base-v2	6.8	5.3	5.4	5.2	5.2	5.3	5.3
all-MiniLM-L6-v2	3.2	3.1	3.0	3.1	3.1	3.1	3.1
paraphrase-multilingual-mpnet-base-v2	4.8	5.0	4.9	5.0	4.8	4.9	4.8
bge-base-en-v1.5	5.7	4.8	4.7	4.8	4.7	4.8	4.7
all-roberta-large-v1	10.5	8.9	8.8	8.7	8.6	8.6	50.1
all-distilroberta-v1	3.0	3.1	3.0	3.1	3.1	3.0	22.6
LaBSE	6.0	5.0	4.8	4.7	4.8	4.8	45.7
paraphrase-MiniLM-L3-v2	2.4	2.1	2.2	2.2	2.2	2.2	3.0
bge-large-en-v1.5	10.5	9.1	8.8	8.6	8.5	8.7	187.2

rises to 50.1 ms for the 5289-record dataset, compared to approximately 8.6 ms for smaller subsets. Similarly, *LaBSE* increases to 45.7 ms for the most extensive dataset. *bge-large-en-v1.5* displays the most significant increase, reaching 187.2 ms for the largest dataset. Conversely, *all-MiniLM-L6-v2* and *paraphrase-MiniLM-L3-v2* demonstrate the most stable processing times across different dataset sizes, indicating more efficient scaling. *all-distilroberta-v1* also shows a slight increase, reaching 22.6 ms for the largest dataset. These results indicate that the computational load associated with processing the complete dataset introduces additional overhead for specific models, revealing differences in how effectively they manage larger inputs.

Table 2 shows the performance of each model in terms of embedding time. The *paraphrase-MiniLM-L3-v2* model demonstrates the most efficient performance, consistently achieving the fastest embedding times across all dataset sizes, with times around 2 ms. In contrast, the *bge-large-en-v1.5* model is the slowest, with embedding times growing with the dataset size and reaching 187.2 ms for the largest dataset of 5289 records. This significant rise suggests that its computational demands scale more substantially with larger inputs. Several other models also show relatively fast performance across the smaller dataset sizes, with times generally below 6 ms, including *all-mpnet-base-v2*, *paraphrase-multilingual-mpnet-base-v2*, and *bge-base-en-v1.5*. However, their embedding times display a noticeable climb for the largest dataset. Specifically, *all-roberta-large-v1* and *LaBSE* attain a moderate expansion, reaching 50.1 ms and 45.7 ms, respectively, for the 5289-record dataset. *all-distilroberta-v1* also shows a jump at the largest dataset size, reaching 22.6 ms.

Most models exhibit relatively consistent embedding times across varying dataset sizes. This suggests that the processing demand remains stable for these models. However, some models experience increased processing duration when handling the largest dataset of 5289 records. For instance, the embedding time for *all-roberta-large-v1* rises to 50.1 ms for the 5289-record dataset, compared to approximately 8.6 ms for smaller subsets. Similarly, *LaBSE* increases to 45.7 ms for the most extensive dataset. *bge-large-en-v1.5* displays the most significant increase, reaching 187.2 ms for the largest dataset. Conversely, *all-MiniLM-L6-v2* and *paraphrase-MiniLM-L3-v2* demonstrate the most stable processing times across different dataset sizes, indicating more efficient scaling. *all-*

distilroberta-v1 also shows a slight increase, reaching 22.6 ms for the largest dataset. These results indicate that the computational load associated with processing the complete dataset introduces additional overhead for specific models, revealing differences in how effectively they manage larger inputs.

A notable similarity across both tables is the consistent performance of specific models, particularly *paraphrase-MiniLM-L3-v2* and *all-MiniLM-L6-v2*. These models yield rapid embedding times for smaller datasets, indicating their efficiency in processing text when the input size is relatively constrained. Furthermore, the models that present the most substantial increase in processing time when handling the larger input are also consistent across both tables, namely *bge-large-en-v1.5*. Regardless of the specific dataset, nearly all models demonstrate an increase in processing time when handling the most significant input, suggesting that scaling to accommodate large datasets introduces computational challenges for these embedding methods.

However, key differences in model's performance were observed between the two tables. The magnitude of the increase in processing time for the most extensive datasets varies considerably. For example, *bge-large-en-v1.5* shows a substantial jump in processing time in both cases, highlighting its less efficient scaling. In contrast, *all-MiniLM-L6-v2* and *paraphrase-MiniLM-L3-v2* maintain more stable and lower processing times even for the largest datasets, indicating better scalability. Additionally, the relative ranking of some models shifts between the tables, with models like *all-roberta-large-v1* and *LaBSE* indicating a more pronounced performance decrease for the largest dataset in Table 1 compared to Table 2, suggesting that dataset characteristics can influence model efficiency.

To investigate a potential correlation with observed increases in processing times near the end of the dataset, Figure 2 reveals sentence length distributions between the two datasets. The "mteb/stsbenchmark-sts" dataset exhibits a broader range of sentence lengths for both the first and second sentences. The distribution is concentrated towards shorter lengths, with two peaks: in the middle and another near the end, indicating the presence of some longer sentences. In contrast, the "sts-companion" dataset shows a strong bias towards very short lengths for the first sentences, while the second sentences, though also predominantly short, display greater variability and include some exceptionally long sentences.

Despite these differences in sentence length distribution, it is unlikely that sentence length alone is the primary driver of the observed increase in processing time for the larger datasets in Tables 1 and 2. While transformer-based models can show increased computational cost with more extended sequences, the relatively modest differences in average sentence length, especially when considering the substantial jump in processing time for some models with the complete datasets, suggest that factors such as batching overhead, memory limitations, or model architecture are more influential in the observed performance scaling.

Figure 3 presents a comparative analysis of various sentence embedding models across different evaluation metrics as the number of training records increases. Generally, for smaller subsets of the mteb-stsbenchmark-sts dataset, the performance of most models appears relatively stable across all evaluation metrics. However, as the dataset size increases towards the full 8628 records, the distinctions between model performances become more apparent. For the error rates (MSE, RMSE, and MAE), some models reveal a degradation in effectiveness, showing increased error with larger datasets. Interestingly, a few models seem to maintain or even slightly improve their error metrics as the data size grows. In contrast to the varied trends in error metrics, Spearman's correlation coefficient and Kendall's Tau tend to follow a somewhat similar overall trend across most models: performance generally increases and then plateaus or slightly decreases as the dataset size expands, suggesting a common pattern in how well these models capture semantic ranking with more data.

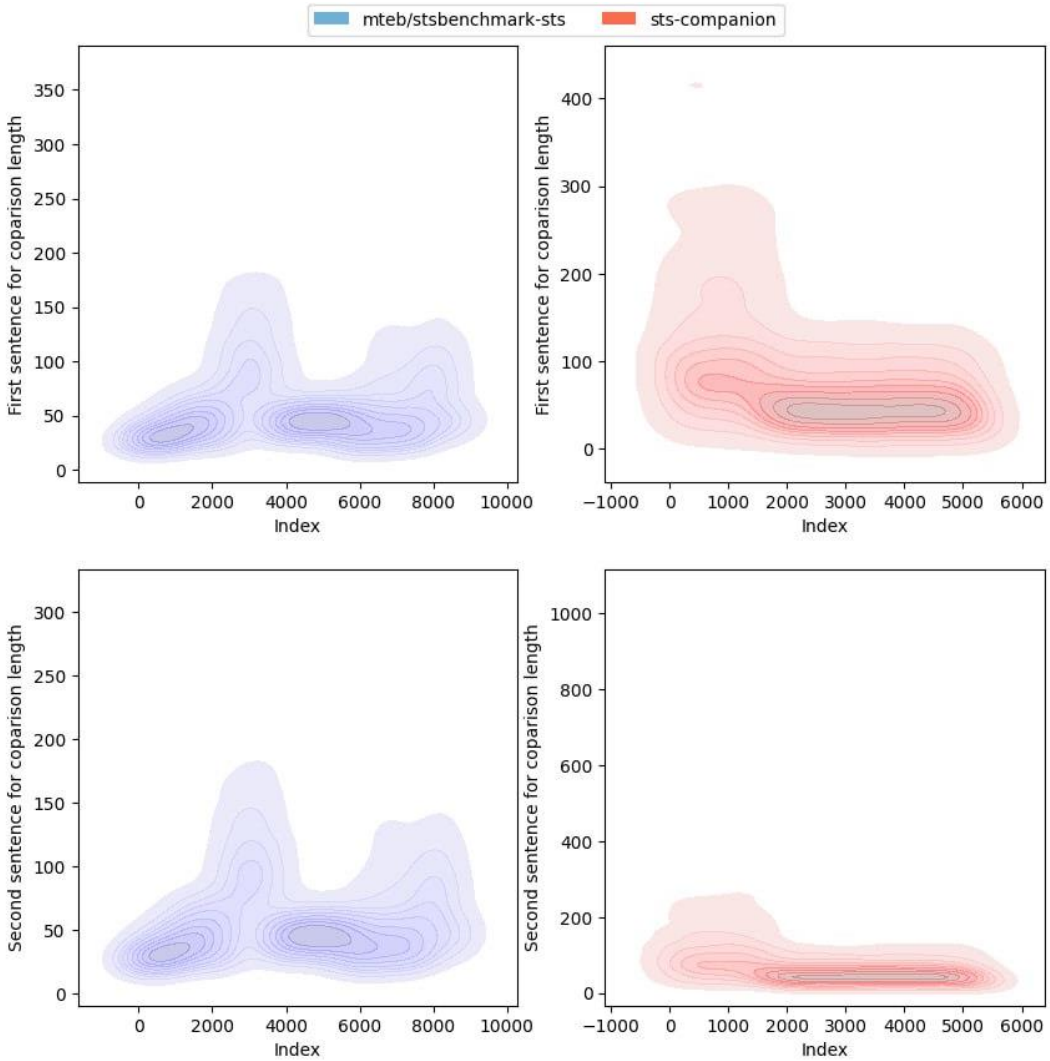


Fig. 2. Sentence lengths in mteb-stsbenchmark-sts and sts-companion.

The *sentence-transformers/all-roberta-large-v1* model appears to be the top performer, consistently demonstrating high R^2 values and low error rates, indicating a strong ability to model the underlying data. Following closely is *BAAI/bge-base-en-v1.5*, which also shows competitive results in terms of R^2 and error metrics, and excels in capturing data relationships as evidenced by its high Spearman's and Kendall's Tau correlation coefficients. The *all-mpnet-base-v2* model generally occupies the third position, showing stable and competitive performance across most metrics. In contrast, the *sentence-transformers/paraphrase-MiniLM-L3-v2* model consistently lags behind the others, showing the lowest R^2 values and the highest error rates throughout the evaluated data range, suggesting it may be less effective in capturing the semantic nuances within this dataset.

Based on the evaluation metrics presented in Figure 3, the *sentence-transformers-/LaBSE* model consistently demonstrates the weakest performance across a range of dataset sizes. Evidenced by the highest MSE and RMSE, *sentence-transformers/LaBSE* attains the largest prediction errors compared to other models. Furthermore, its MAE also remains notably high. In terms of correlation, the *sentence-transformers/LaBSE* model

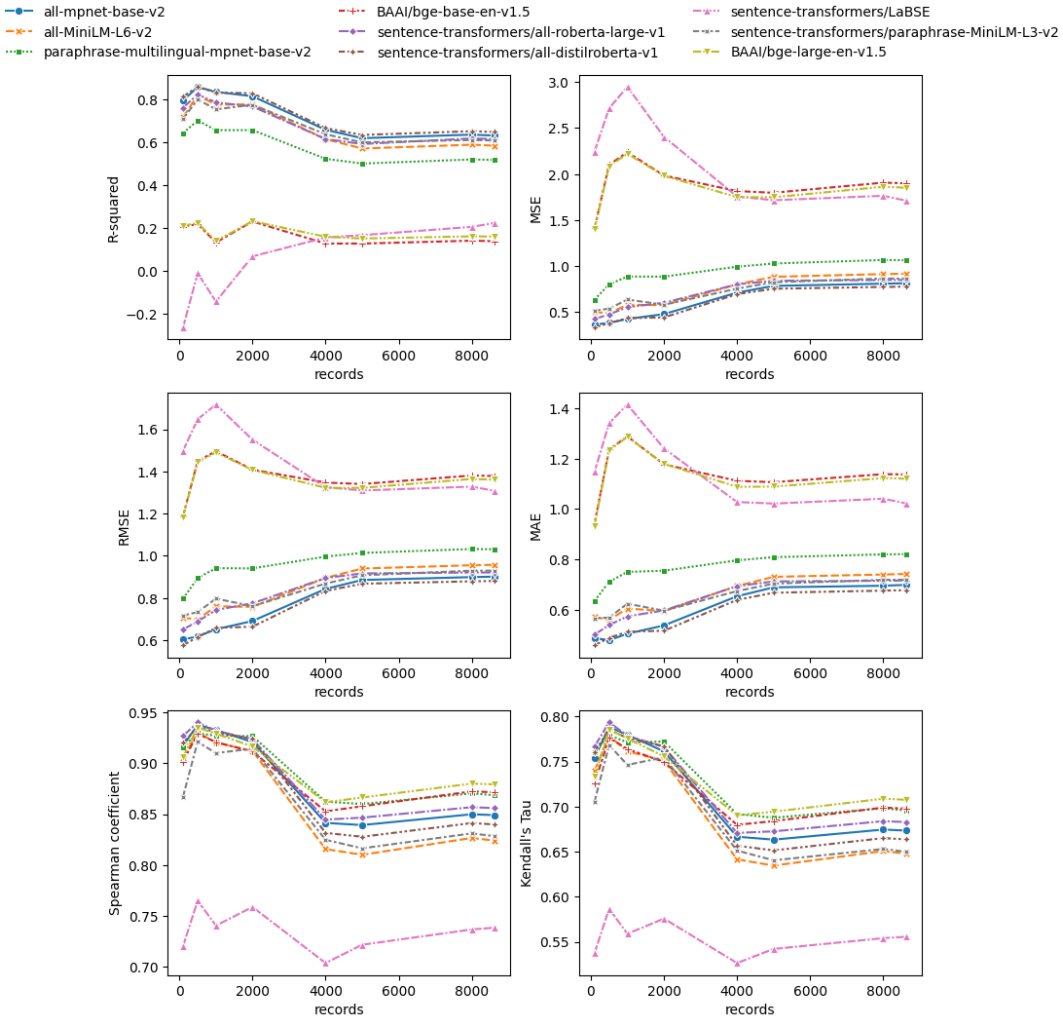


Fig. 3. Performance metrics for sentence embedding models across varying mteb-stsbenchmark-sts dataset sizes.

yields the lowest Spearman's and Kendall's Tau coefficients, indicating a poor ability to capture both the monotonic and ordinal relationships between predicted and actual semantic similarities. While its R^2 values show some improvement with larger datasets, its overall performance across the majority of the key evaluation metrics positions *sentence-transformers/LaBSE* as the least effective model among those compared in this analysis.

Considering the time taken for vector representation in Table 1 alongside the performance metrics in Figure 3 reveals interesting trade-offs. The *all-MiniLM-L6-v2* model stands out as remarkably fast, consistently achieving the lowest processing times across all dataset sizes. However, its performance, while generally competitive, doesn't place it at the top tier, often showing lower R^2 and higher error rates compared to models like *all-roberta-large-v1* and *bge-base-en-v1.5*. Conversely, models like *sentence-transformers/all-roberta-large-v1* and *BAAI/bge-large-en-v1.5*, which demonstrated strong performance in the evaluation metrics, tend to be significantly slower, particularly as the dataset size increases, as seen in the higher time values in Table 1. The *sentence-transformers/paraphrase-MiniLM-L3-v2*, while being relatively fast, also shows the weakest performance

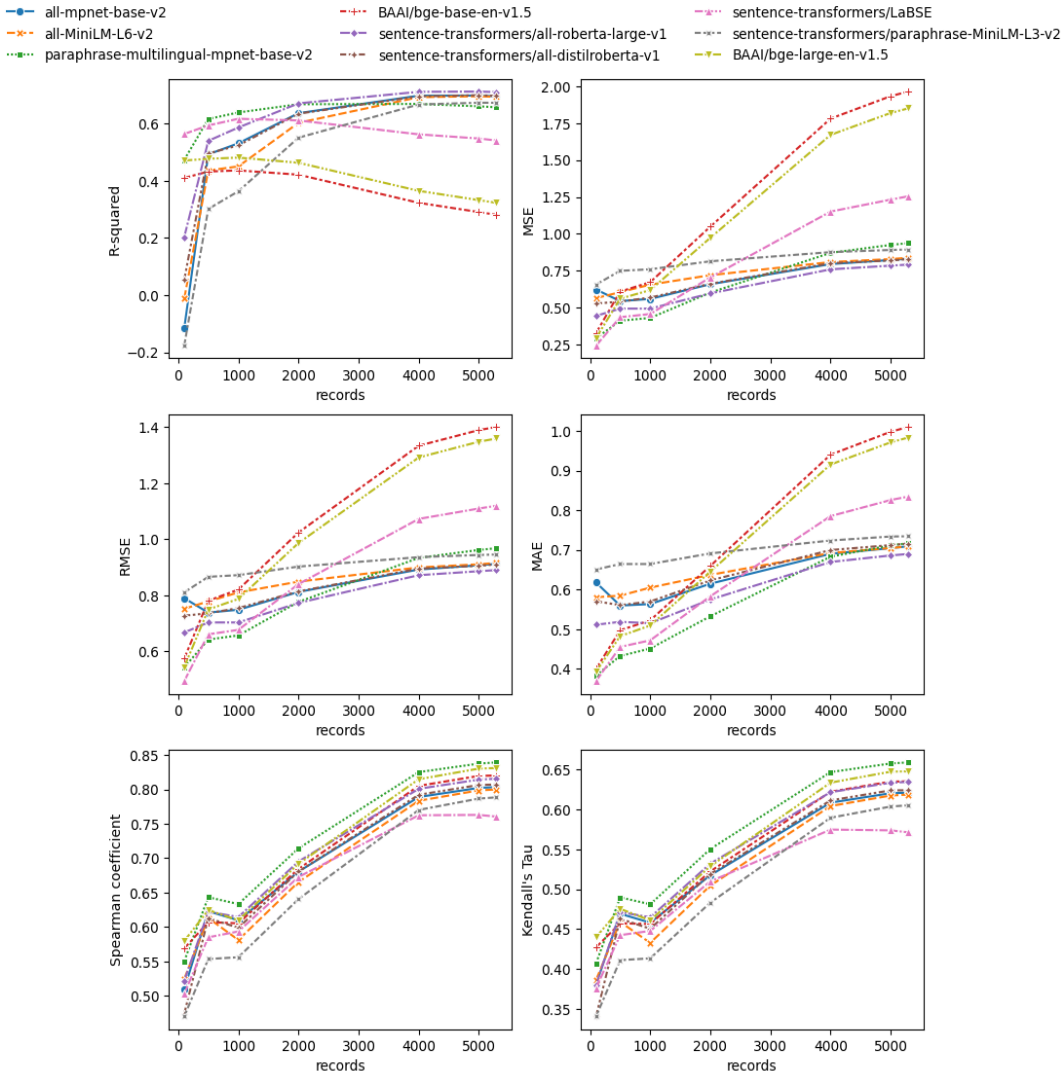


Fig. 4. Performance metrics for sentence embedding models across varying sts-companion dataset sizes.

across the metrics, making it a less appealing choice when considering both efficiency and accuracy.

Figure 4 presents a performance comparison of various sentence embedding models across several evaluation metrics as the number of training records increases. Generally, the models show the most significant performance changes within the initial range of data points, tending to plateau beyond approximately 1000-2000 records. Across the metrics, distinct performance characteristics emerge for different models. The *sentence-transformers/all-roberta-large-v1* model stands out for attaining high R^2 values and low error rates, indicating a strong ability to fit the data and produce accurate predictions. However, when considering the correlation metrics, *paraphrase-mulrilingual-mpnet-base-v2* demonstrates superior performance, consistently presenting the highest Spearman's and Kendall's Tau coefficients, suggesting a stronger ability to capture the monotonic and ordinal relationships within the data. In contrast, the *BAAI/bge-base-en-v1.5* model consistently exhibits the lowest R^2 values and the highest error rates, along with generally lower correlation coefficients, indicating it is the least effective in capturing the underlying

relationships within this dataset based on these metrics. Therefore, the choice of the "best" model depends on the specific priorities of the application, whether it emphasizes overall prediction accuracy or the preservation of data rankings.

Based on the evaluation metrics in Figure 4, the top-performing models present a trade-off. For achieving high prediction accuracy, as indicated by R^2 and error metrics, *sentence-transformers/all-roberta-large-v1* appears to be the most effective. However, for tasks where the preservation of data relationships is paramount, the *paraphrase-mulrilingual-mpnet-base-v2* model demonstrates superior performance in terms of Spearman's and Kendall's Tau coefficients. Consequently, the choice between these two as the "best" depends on the specific application priorities. The *all-mpnet-base-v2* model generally is the third-best option, showing a consistent and competitive performance across all the evaluated metrics.

Focusing solely on Figure 4, the *BAAI/bge-base-en-v1.5* model demonstrates the weakest performance in terms of prediction accuracy, showing the lowest R^2 values and the highest error rates. When considering the correlation metrics, the situation changes with increasing data. Initially, the *sentence-transformers/paraphrase-MiniLM-L3-v2* model generally shows lower Spearman's and Kendall's Tau values. However, as the number of records increases, the performance of *LaBSE* deteriorates in terms of both Spearman's and Kendall's Tau, eventually becoming the lowest among all models in the higher record range. This suggests that while *BAAI/bge-base-en-v1.5* struggles most with prediction accuracy, *LaBSE's* ability to preserve the relative ordering of the data appears to degrade with larger datasets within this evaluation, ultimately making it the least effective in this aspect at higher record counts.

Combining the performance results from Figure 4 with the timing data from Table 2 reveals a trade-off between model accuracy and computational efficiency. The *all-MiniLM-L6-v2* model, which exhibits the fastest processing times, demonstrates reasonably competitive performance, though it doesn't excel in any single metric. In contrast, while *sentence-transformers/all-roberta-large-v1* shows strong accuracy, it is considerably slower, as indicated by its higher processing times. The *BAAI/bge-large-en-v1.5* model, which performs well in capturing data relationships, also becomes notably more time-consuming as the dataset size increases.

Comparing the performance of the sentence embedding models across the two datasets (mteb-stsbenchmark-sts in Figure 3 and sts-companion in Figure 4) reveals both consistent patterns and interesting divergences. Notably, *sentence-transformers/all-roberta-large-v1* consistently demonstrates strong performance in terms of prediction accuracy on both datasets, achieving high R^2 values and low error rates. Similarly, *all-mpnet-base-v2* maintains a relatively stable and competitive performance across the board in both evaluations. This suggests that these models possess a degree of robustness and generalizability across different semantic textual similarity tasks. Furthermore, the tendency for performance gains to diminish with increasing data size (beyond a certain point) is observed in the results for both datasets, indicating a potential saturation point for these model architectures on this type of task.

However, the identification of the weakest performing model differs between the two evaluations. In Figure 3, *sentence-transformers/paraphrase-MiniLM-L3-v2* generally shows the poorest performance across most metrics. In contrast, Figure 4 highlights *BAAI/bge-base-en-v1.5* as having the lowest accuracy and reveals a degradation in *LaBSE's* correlation performance with larger data. This discrepancy suggests that the specific characteristics of each dataset can influence the relative strengths and weaknesses of the models. The sts-companion dataset, for instance, might present different semantic nuances or complexities that impact the models in varying ways compared to the *mteb-stsbenchmark-sts* dataset. Therefore, while some models have consistent behavior, the optimal choice depends on the utilized dataset.

The comparative performance and efficiency data detailed in this section can be used as a practical guide for model selection, enabling practitioners to weigh crucial trade-offs between predictive effectiveness, nuanced semantic understanding, and computational cost. This allows for the identification of the most suitable model tailored to the specific task requirements and available computational resources within an English-language context.

CONCLUSION

This study provided a comparative analysis of several sentence embedding models, evaluating their performance on semantic similarity tasks across two datasets and assessing their embedding generation efficiency. The findings highlight significant trade-offs between model accuracy and speed, alongside model-specific strengths and weaknesses in capturing semantic relationships. The results offer insights for selecting the most suitable model based on application-specific priorities and resource constraints.

In terms of performance metrics, *all-roberta-large-v1* consistently demonstrated strong results in prediction accuracy, achieving high R^2 values and low error rates across both datasets. For capturing semantic relationships and preserving the relative ordering of sentences, *BAAI/bge-large-en-v1.5* generally excelled, particularly evident in its high Spearman's and Kendall's Tau coefficients. *all-mpnet-base-v2* consistently offered a balanced and competitive performance across all metrics.

Conversely, the models yielding the weakest performance varied depending on the metric and the dataset. *BAAI/bge-base-en-v1.5* tended to show the lowest R^2 and highest error rates on the sts-companion dataset. Regarding the preservation of nuanced semantic understanding, *sentence-transformers/paraphrase-MiniLM-L3-v2* generally showed lower correlation values, while *LaBSE's* performance in Spearman's and Kendall's Tau deteriorated significantly with larger datasets in the "sts-companion" evaluation.

Considering the interplay between performance and computational efficiency, *sentence-transformers/all-MiniLM-L6-v2* model stands out as the most time-efficient, consistently generating embeddings rapidly across varying dataset sizes, albeit with a slight compromise in top-tier semantic representation accuracy. In contrast, models demonstrating high accuracy, such as *all-roberta-large-v1*, often incurred a higher computational cost. Notably, *sentence-transformers/LaBSE* generally exhibited weaker performance metrics coupled with moderate to slower processing speeds, making it a less favorable choice when considering both factors.

Moving forward, we plan to apply the top-performing models from this evaluation, such as *all-mpnet-base-v2*, *sentence-transformers/all-distilroberta-v1*, and *sentence-transformers/all-MiniLM-L6-v2*, to custom datasets. This will allow us to assess their generalization capabilities and fine-tune them for specific applications. Future work may also explore techniques to optimize these models for efficiency and performance on domain-specific data.

AUTHOR CONTRIBUTIONS

Conceptualization, [*M.S.*]; methodology, [*M.S.*]; validation, [*M.S.*]; writing – original draft preparation, [*M.S.*]; writing – review and editing, [*B.P.*, *M.S.*]; supervision, [*B.P.*].

All authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] Bengio, Y., Ducharme, R., Vincent, P., & Jauvin, C. (2003). A neural probabilistic language model. *Journal of machine learning research*, *3*(Feb), 1137-1155. https://doi.org/10.1162/153244303322533223

[2] Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. https://doi.org/10.48550/arXiv.1301.3781

[3] Pennington, J., Socher, R., & Manning, C. D. (2014, October). Glove: Global vectors for word representation. In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP) (pp. 1532-1543). https://doi.org/10.3115/v1/D14-1162
[4] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30. https://doi.org/10.48550/arXiv.1706.03762
[5] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019, June). Bert: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers) (pp. 4171-4186). https://doi.org/10.18653/v1/N19-1423
[6] Pavlyshenko, B., & Stasiuk, M. (2024). Data augmentation in text classification with multiple categories. Electronics and information technologies, (25). http://dx.doi.org/10.30970/eli.25.6
[7] Pavlyshenko, B., & Stasiuk, M. (2025). Using Large Language Models for Data Augmentation in Text Classification Models. International Journal of Computing, 24(1), 148-154. https://doi.org/10.47839/ijc.24.1.3886
[8] Pavlyshenko, B. (2014). Clustering of authors' texts of english fiction in the vector space of semantic fields. Cybernetics and Information Technologies, 14(3), 25-36. https://doi.org/10.2478/cait-2014-0030
[9] Pavlyshenko, B. (2013). Classification analysis of authorship fiction texts in the space of semantic fields. Journal of Quantitative Linguistics, 20(3), 218–226. https://doi.org/10.1080/09296174.2013.799914
[10] Han, M., Zhang, X., Yuan, X., Jiang, J., Yun, W., & Gao, C. (2021). A survey on the techniques, applications, and performance of short text semantic similarity. Concurrency and Computation: Practice and Experience, 33(5), e5971. https://doi.org/10.1002/cpe.5971
[11] Risch, J., Möller, T., Gutsch, J., & Pietsch, M. (2021). Semantic answer similarity for evaluating question answering models. https://doi.org/10.48550/arXiv.2108.06130
[12] Vrbanec, T., & Meštrović, A. (2017, May). The struggle with academic plagiarism: Approaches based on semantic similarity. In 2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO) (pp. 870-875). IEEE. https://doi.org/10.23919/MIPRO.2017.7973544
[13] Cer, D., Diab, M., Agirre, E., Lopez-Gazpio, I., & Specia, L. (2017). Semeval-2017 task 1: Semantic textual similarity-multilingual and cross-lingual focused evaluation. https://doi.org/10.18653/v1/S17-2001
[14] mteb/stsbenchmark-sts. Retrieved from https://huggingface.co/datasets/mteb/stsbenchmark-sts
[15] Song, K., Tan, X., Qin, T., Lu, J., & Liu, T. Y. (2020). Mpnet: Masked and permuted pre-training for language understanding. *Advances in neural information processing systems*, *33*, 16857-16867. https://doi.org/10.48550/arXiv.2004.09297
[16] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach. https://doi.org/10.48550/arXiv.1907.11692
[17] Feng, F., Yang, Y., Cer, D., Arivazhagan, N., & Wang, W. (2020). Language-agnostic BERT sentence embedding. https://doi.org/10.48550/arXiv.2007.01852

АНАЛІЗ СЕМАНТИЧНОЇ СХОЖОСТІ З ВИКОРИСТАННЯМ ВЕКТОРНИХ ПРЕДСТАВЛЕНЬ РЕЧЕНЬ НА ОСНОВІ ТРАНСФОРМЕРІВ

Богдан Павлишенко **, *Микола Стасюк*** ***
Львівський національний університет імені Івана Франка,
вул. Драгоманова 50, 79005 Львів, Україна

АНОТАЦІЯ

Вступ. Моделі трансформерів стали центральним елементом обробки природної мови, демонструючи еталонну продуктивність в оцінці семантичної схожості, що є критично важливим завданням для різноманітних додатків. Ці моделі фіксують деталі взаємозв'язків між текстами, розширюючи можливості оцінки семантичної спорідненості.

Матеріали та методи. Ефективність моделей вбудовування речень, зокрема all-mpnet-base-v2, all-MiniLM-L6-v2, paraphrase-multilingual-mpnet-base-v2, bge-base-en-v1.5, all-roberta-large-v1, all-distilroberta-v1, LaBSE, paraphrase-MiniLM-L3-v2, bge-large-en-v1.5, на двох наборах даних. До даних застосовано такі кроки попередньої обробки: переведення в нижній регістр, видалення стоп-слів, очищення від спеціальних символів і цифр та лематизація. Від'ємні значення косинусного коефіцієнта подібності, що вказують на семантичну несхожість, розглядалися як еквівалент оцінки подібності, що дорівнює 0, а невід'ємні значення косинусного коефіцієнта подібності масштабувалися до діапазону 0-5. Для оцінки використовувалися такі показники, як R^2, MSE, RMSE, MAE, коефіцієнт кореляції Спірмана та коефіцієнт кореляції рангу Кендала.

Результати. Продуктивність моделей загалом покращується зі збільшенням обсягу даних. За результати оцінювання моделей вбудовування речень виявилено різницю в ефективності. all-roberta-large-v1 продемонструвала високу точність з високими значеннями R^2 і низькими помилками. BAAI/bge-large-en-v1.5 чудово вловлює семантичні зв'язки, демонструючи високі значення коефіцієнтів Спірмена та Кендалла. all-MiniLM-L6-v2 демонструє найшвидше генерування вбудовувань. BAAI/bge-base-en-v1.5 показує найнижчу точність. Час обробки, як правило, збільшується зі збільшенням розміру даних.

Висновки. Це дослідження висвітлює компроміс між продуктивністю та обчислювальною ефективністю при вбудовуванні речень. Вибір моделі залежить від балансування цих факторів відповідно до конкретних потреб програми. У випадках, коли вимагається висока точність, слід надавати перевагу all-roberta-large-v1, а тоді як пріоритетом є швидкість, краще використовувати all-MiniLM-L6-v2. BAAI/bge-large-en-v1.5 найкраще підходить для завдань, що вимагають розуміння семантичних деталей.

Ключові слова: семантична схожість, вбудування речень, трансформери.

Received / Одержано	Revised / Доопрацьовано	Accepted / Прийнято	Published / Опубліковано
15 May, 2025	02 June, 2025	09 June, 2025	26 June, 2025

Electronics and Information Technologies, 2025, 30, 59–74
http://publications.lnu.edu.ua/collections/index.php/electronics/index

UDC 004.9

INFORMATION TECHNOLOGY FOR DATA AUTHENTICATION BASED ON BLOCKCHAIN TITLE OF MANUSCRIPT

***Victoria Vysotska*[1], *Oleg Prokipchuk*[1], *Mariia Nazarkevych*[2]*, *Roman Romanchuk*[1]**

[1]*Department of Information Systems and Networks, Lviv Polytechnic National University 12 Stepana Bandery St., Lviv, 79013, Ukraine*

[2]*Department of Radiophysics and Computer Technologies, Ivan Franko National University of Lviv 50 Drahomanova St., Lviv, 79005, Ukraine*

Vysotska, V.A., Prokipchuk, O.A., Nazarkevych, M.A., Romanchuk, R.V. (2025). Information Technology for Data Authentication Based on Blockchain Title of Manuscript. *Electronics and Information Technologies*, 30, 59–74. https://doi.org/10.30970/eli.30.5

ABSTRACT

Background. The article describes the process of developing a blockchain-based technology for information support of the processes of authentication of goods. The paper also investigates the performance of the developed system depending on the adjustment of four parameters: mining complexity, dependence of mining time on technical support capacity, dependence of block creation time on the number of transactions, and dependence of blockchain validation time on the number of blocks.

Materials and Methods. The mining complexity was analyzed in the range from 1 to 7 (with a complexity of 8 or more, mining on a working device takes too long). This parameter determines the number of zeros that must be at the beginning of the hash to consider the work as confirmed. Each subsequent difficulty increases the total time several times. This is due to a significant increase in the number of operations that need to be performed during mining. Of course, mining also depends on the capacity of the technical support, which should be a complex of advanced processors and video cards. The analysis of the dependence of mining time on the processor frequency was conducted for a mining complexity of 5.

Results and Discussion. Additional processor power can significantly reduce mining time. The next study is to analyze the dependence of block creation time on the number of transactions. The main resource burden of this stage is the calculation of hash functions for transactions and for the block, as well as the construction of the hash tree. From the data obtained, we can conclude that the number of transactions in a block is its main resource load. This should be taken into account when choosing the maximum number of transactions per block. The last study was to identify the dependence of blockchain validation time on the number of blocks. The data shows a linear dependence of the blockchain validation time on the number of blocks.

Conclusion. This indicates that the validation process is not overloaded with resource-dependent operations. As the blockchain expands over time, the expected duration of blockchain validation can be calculated according to the following linear relationship.

Keywords: network communication, information system, counterfeit goods, blockchain technology, product life cycle, intelligent information retrieval system

INTRODUCTION

The fight against counterfeiting is ongoing every day, and many methods of identifying authentic goods have already been developed [1-3]. The most popular are the legislative method, the method of developing recommendations, the method of making it difficult to copy, the method of monitoring suspicious ads, and the method of keeping records of each unit of goods [4-6]. The system under development belongs to the latter category. The market for the use of anti-counterfeiting methods is quite developed, as most companies already use at least one of these methods. Among similar systems, Microsoft Aura Ledger stands out, which also uses blockchain technology. It is delivered in the form of software for producers of goods. The advantage of the product under development is that it unites many manufacturers into a single network to ensure greater reliability and transparency of data, as well as the use of two-stage identification, which provides greater accuracy. The main consumers are manufacturers of goods that suffer material damage from counterfeiters [7-11]. Such producers want to minimize the impact of counterfeit goods on their profits, as well as to maintain the reputation and trust of their customers. Competitors include other anti-counterfeiting software providers with the same level of protection.

First, it is necessary to analyze the available tools for solving the problem and select those that are best suited for use in implementing the information system. In order to form a set of tools, first of all, it is necessary to define the tasks to be solved. The main tasks are as follows:

- The developed software needs to be created that allows data and blockchain to be operated in dialogue mode with the user interface.
- The developed software should have high performance of arithmetic and algorithmic operations, be flexible and be able to be customized and scaled.
- The developed software should be able to connect to other applications, in particular to the same software object on other end devices, to combine them into a P2P network.
- Need to create a P2P server to manage information exchange between nodes.
- Need to create a client application in the form of a mobile app or website.

MATERIALS AND METHODS

Once the list of tasks has been compiled, the selection of tools and technologies for their implementation begins. First of all, technical means for implementing the main software application are selected. The technologies used in the creation of the program are as follows: Java, Spring Boot, JavaFX, GSON, and Maven. The same set of technologies was used to implement the P2P server and the server part of the website. The following technology stack was used to create the client part: HTML, CSS, JavaScript. First of all, it is necessary to choose the main programming language. Among the possible options are: Java, Python, C++, C#. To achieve high performance, the language must be compiled. From the available options, the Java programming language was chosen. Java is an object-oriented programming language developed by Sun Microsystems, which was later acquired by Oracle. Programs created using this language are compiled into bytecode, which allows for high performance. Java is compiled into a special code that is recognized by the Java Virtual Machine (JVM). This allows programs to run on any operating system that supports JVM. This multi-platform capability is a big plus for the system being developed, as it allows clients and servers to be developed for different platforms [12]. Java is also object-oriented, which allows the system being developed to be flexible and scalable [13]. The first of the necessary tools should be a framework for centralized object management. Modern enterprise applications are very large and consist of hundreds of classes. With this kind of organization, an application can quickly become highly coupled and non-scalable. To avoid this, the application must adhere to the following rules [14]:

- Application classes should aim for Low Coupling.
- Each class should have only one duty (Single Responsibility Principle).

- When one object creates another object, it performs the duty of creating objects, i.e., according to SRP, it should not do anything else but construct objects. Also, very often objects perform several functions at once, even without considering the creation of other objects.

There are 3 design patterns to help in this situation: strategy, control inversion, and dependency injection.

- Strategy is a design pattern that allows a class to extend its functionality by delegating additional work to auxiliary objects. Thanks to this pattern, all classes will have one main duty and delegate all others.
- Dependency injection is a design pattern for changing the order in which auxiliary objects are assigned to the main object. The class does not create auxiliary objects itself but only declares containers for them. The controlling program injects the auxiliary objects into the main one. In this way, the object avoids the obligation to construct objects.
- Control Inversion is a design pattern that is a module that registers and constructs application objects. This pattern allows you to automate the process of dependency injection, thereby significantly reducing the complexity of the program [15].

The use of these design patterns allows for high flexibility and scalability of the application. For the system under development, 2 possible frameworks were chosen: Google Guice and Spring Boot. I chose the latter because of its greater functionality [16]. The next step is to choose a tool for creating the user interface. The choice of such tools for the Java programming language is small and consists of two tools: Swing and JavaFX [17].

Both tools allow the creation of high-quality user interfaces and were developed under the leadership of Oracle. Starting with Java version 9, JavaFX is no longer included in the core JDK development package and is being developed separately from Oracle. JavaFX was chosen for this project because of its convenience and superior functionality, in particular the ability to create designs in FXML files with XML markup, whereas in Swing, components can only be created and populated programmatically. One of the processes that requires a lot of extra time to implement manually is serialization and deserialization. Serialization is the conversion of an object into a byte or character format that is convenient for transfer to non-Java environments. Deserialization is the reverse process, i.e. converting a character or byte format back into a software object. For this program, the JavaScript Object Notation (JSON) character data transfer format was chosen because it is simple to implement and easy to read. This is necessary for tracking and demonstrating the intermediate results of the programme's work. Possible implementations of this technology are Google GSON [19] and Jackson. GSON was chosen because of its higher popularity, which ensures more active work on the software product.

For convenient development and ensuring the correct interaction of all software components, it is advisable to use one of the building systems. These systems automatically compile and link all modules into a single software package and provide the ability to control the versions of all components and the main product. There are 3 such systems for Java: Ant, Maven, and Gradle. Today, Ant is an outdated system. Of the two modern systems, Maven [20] was chosen because of the widespread use of its repositories. Although Gradle is a more modern and advanced system, not all components are compatible with it yet.

One of the key points in developing the main program is network communications. A method of data exchange using the TCP protocol is required. Unlike UDP, TCP [21] checks sent data for possible losses, which is very necessary in a secure and accurate system. One way to achieve this is to use the HTTP application layer protocol, as it is based on TCP and there are many frameworks for Java that allow this functionality to be easily implemented. Despite this, standard Java functionality was chosen to create Socket/Server connections for all P2P communications. This method allows a channel to be established between the client and the server, both of which can listen and edit. This approach is

necessary for implementing a P2P network because the communication of such a connection is not regulated by higher-level protocols and can therefore be arbitrary for the developed product. Although the HTTP protocol was rejected for the implementation of a P2P network, it is an integral part of a web server. For the client application, the option of a website that communicates with the server using the HTTP protocol was chosen. HTTP is an application-level protocol based on the TCP/IP protocol stack. Communication between the server and the client is performed using a request-response pattern. The client forms a request, fills it with data and sends it to the server. In turn, the server processes the request and sends a response to the client. Communication takes place using one of the following methods: GET, POST, PUT, DELETE, OPTIONS, HEAD, PATCH, TRACE, CONNECT [22]. In most cases, the first four methods are used. There are several basic ways to implement such a server, but they are all based on the use of a Servlet container. A servlet is a software unit (object) that processes a client request in a separate thread. The most popular implementation of a Servlet container is Tomcat [23]. However, Spring Boot, which is already used in the program under development, contains extensions compared to Tomcat, which increase its functionality and simplify interaction with other Spring Boot components, which is why this tool is chosen for the implementation of the web server.

In turn, the client side is a web application. Any website is based on three main technologies: HTML, CSS, and JavaScript. This technology stack has no current analogues, so the choice is quite simple. At the beginning of the Internet's development, various tools could be used to create a website, but over time, they have all been ousted from the market.

JavaScript is a scripting language that is executed every time a website is launched. It is used to execute all the logic of a website, as well as to provide user interaction with the website and to asynchronously exchange data with the server. The language's runtime environment, such as a browser, usually restricts the language's access to the resources of the owner's device, so visiting websites is completely safe, as the execution of dangerous scripts has been prevented. Since the language is scripted, the user could call new commands directly while using the website [24]. When the technology stack has already been formed, it is time to choose software development and testing environments. The first such environment should be a tool for developing, compiling, and running Java code. Today, the main competitors on the market are Eclipse and JetBrains Intellij Idea. NetBeans used to be an active participant in this competition, but now its market share is very small. Based on a general review of previous products, it can be said that these are very powerful and functional tools. A more detailed comparison is shown in Table 1. To perform this work, we obtained a student license for the Intellij Idea Ultimate Edition [25].

The second necessary tool is a means of debugging and testing the client web application. Most browsers today have built-in tools for this purpose. The most popular among them are Google Chrome, Firefox, and Opera. Among the above-listed options, Google Chrome was chosen due to its functionality and ease of use. All hardware and software tools are used to perform this work in a specific configuration. Each of them contains its own unique parameters that characterize a particular application. Most tools contain such characteristics as product version, developer, system requirements, etc. For selected software products, here are their detailed specifications.

To date, blockchain has become widespread and is used in many areas of everyday life, especially in cryptocurrencies such as Bitcoin, Ethereum, Binance Coin, Tether, Bitcoin cash, Litecoin, and others [26]. The principle of blockchain operation is shown in Fig. 1.

The main qualitative characteristic of this technology is the security and transparency of this data structure, as new blocks can only be added to the blockchain without the possibility of editing them. This means that any data entered into the blockchain becomes visible to all its participants and remains there in its original format for the entire duration of the blockchain's existence. There are many blockchain implementations, so its behavior

Table 1. **Comparison Intellij Idea and Eclipse**

Parameters	Intellij Idea	Eclipse
System requirements	Minimum 2 GB of RAM	Minimum 0.5 GB of RAM
Distribution method	Free with paid version	Free
Debugging	An extended set of debugging tools	Standard debugging tools
Plugins	750+ plugins	1250+ plugins
Auto-complete	Automaton	Use the key combination Ctrl + Space
Productivity	Optimized for indexed transactions	Faster under heavy loads
Refactoring	An extended set of tools	Standard set of tools
Design	Modern design, easy to use	Outdated and overloaded design
Focus	Small and medium-sized projects	Large projects

and characteristics can often vary from one product to another. The first successful and most well-known blockchain implementation, Bitcoin, was taken as a basis for this paper. A block is the basic structural unit of a blockchain. The only operation to change the blockchain is to add a new block to an existing chain. Blocks are created by the nodes of the blockchain network. In addition to being a structural unit of the chain, the most important function of a block is to be a receptacle for data that must be unchanged. A block consists of a header and basic data. The header of a block is also called its metadata. This division was created due to the fact that the blockchain can reach very large sizes with prolonged use. Thus, the Bitcoin cryptocurrency today occupies more than 200 GB. This is not a problem for specialized desktop devices, but for mobile devices, downloading 200 GB of data can be a challenge. For optimization purposes, the block was divided into a header containing the minimum required data set and the main body. This makes it possible to develop mobile clients that work only with block headers. Such mobile clients have one drawback. If necessary, they have to download the necessary data, and when using such

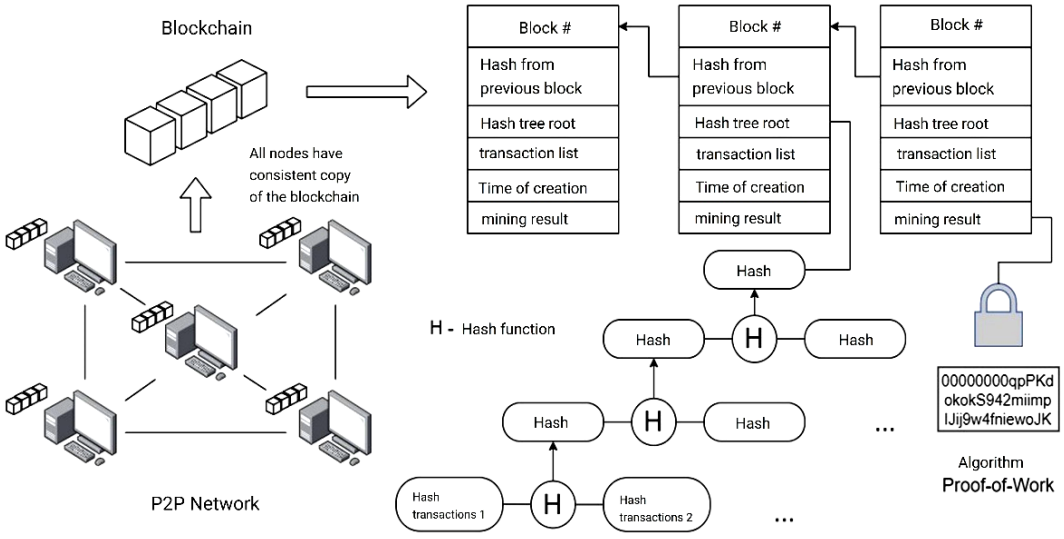


Fig. 1. Principles of blockchain robots.

Table 2. **Block structure**

Field	Size
Caption	
Hash of the previous block	32 Byte
Hash block	32 Byte
Hash root	32 Byte
Hour of creation	8 Bytes
Appendix (nonsense)	8 Bytes
Main part	
Hash tree	Depending on the number of transactions
List of transactions	Limited to implementation

a client, it is very important to make sure that this download is carried out from a reliable source. The structure of the block is shown in Table 2. Thus, the size of the block header is 112 Bytes. Let's take a closer look at each of the components [27].

A Block Hash is a unique block identifier obtained by passing the block header through a hash function. The SHA-256 hashing algorithm was used to perform this work. The algorithm generates 256-bit hash for any data. To form a chain, the blocks must be connected in a certain way. Each block contains a reference to the previous block in the form of a hash of the previous block (Previous Block Hash). As a result, having the last block, it is possible to trace the initial block by moving through the previous hashes. The hash of the previous block is included in the data used to calculate the hash of the current block. This provides protection against data tampering. If the data of any block of the blockchain is changed, all subsequent blocks become invalid. The Merkle root is the root of the hash tree. A Merkle tree is a tree built based on hash values of transactions. The tree is built using the following algorithm:

- The construction starts with the tree leaves, which are the hashes of each of the submitted transactions.
- Transaction hashes form a queue.
- 2 elements are selected from the queue, and a parent node is created on their basis.
- The parent node is created as a result of passing the concatenation of children's hashes through the hash function.
- The resulting node is added to the queue of the next stage.
- The operation is repeated until the current queue is over, after which the queue of the next stage becomes the current one.
- The general algorithm is repeated until there is only 1 element left in the final queue, which is the root of the tree.

Thus, each parent node is a hash of the hashes of the child nodes [28]. Blockchain nodes always accept only the longest chain of blocks, all others are considered irrelevant and discarded. The blockchain is already protected from data tampering by hashing it, but what prevents an attacker from creating a longer chain of blocks so that other nodes will accept the fake blockchain as the real one? The Proof-of-Work algorithm was developed to prevent such a situation. This algorithm is based on the fact that the creation of each block should be accompanied by certain resources and time costs of the processor. The process that provides such costs is Mining. Mining a block means selecting a Nonce value such that the block's hash starts with a certain number of zeros. The number of zeros that should be at the beginning of the hash is determined by the complexity of mining. Difficulty is a configuration parameter that can be used to control the duration of mining. There are

two main algorithms for selecting a Nonce value: iteration and the use of random numbers. The use of random numbers allows this algorithm to be executed in parallel in multithreaded mode to increase mining performance. Thus, in order to create a fake chain that is longer than the actual one, it is necessary to perform Proof-of-work for each block of the new chain. If the mining time is 10 minutes and an attempt is made to compromise a blockchain with a length of 1,000 blocks, this will take 10,000 minutes. This makes such attacks virtually impossible. The main data of a block is its transactions. A transaction stores secure data that can interact with each other. The person who creates the block (miner) is not necessarily the owner of the transaction. A node receives transactions from other nodes, combines them into a single block, and adds them to the blockchain. The content of a transaction can be anything. In the Bitcoin system and in the application of this paper, the filling is used in the form of transferring a numerical balance from one wallet to another. A transaction consists of: transaction id, sender's address, recipient's address, transferred balance, inputs, outputs, additional data, and an electronic digital signature. Let's take a closer look at each of the elements:

- The transaction Id is its unique identifier assigned after its creation.
- The sender/receiver addresses are the public keys of the wallets.
- A wallet is a set of public and private cryptographic keys. The balance of a wallet can be calculated. A wallet is used by the sender to create a transaction to transfer funds to another wallet. The public key of a wallet is used as its address and to verify digital signatures. The private key of a wallet is used to create a digital signature for a transaction. Public and private keys are generated using the RSA algorithm. The generation of such keys depends on random numbers. In turn, the algorithm for generating such numbers must not have any patterns.
- RSA is an asymmetric cryptographic encryption algorithm based on the properties of large prime numbers. The purpose of the algorithm is to encrypt data and create a digital signature. The algorithm is widespread today and is used in many applications [29].

To start the system, it is necessary to combine the nodes into a single network, as the reliability of the blockchain directly depends on the number of users. When implementing the standard, it was decided to move away from the standard centralized client-server model in favor of the decentralized P2P model. P2P is a network topology in which each participant can simultaneously act as a client, making requests to one group of nodes, and be a server, responding to requests from another group of nodes. Such a network is not hierarchical, and all its endpoints are on the same level. If one of the nodes fails, the network can still operate correctly. The nodes of such a network are called peers [30]. There are several ways to organize such a network: decentralized, centralized, and mixed. Decentralized P2P consists only of nodes located on the 1st level, and neighboring peers are used to find nodes. In a centralized P2P, there is a server that knows the addresses of the nodes, and when establishing communication, peers contact the server to get the addresses of other peers. The mixed method uses both approaches.

In a blockchain, each pir has its own local copy of the blockchain. When a node starts working, it scans the network for the longest chain of blocks, and if it finds one, it updates its local copy. When a node creates a block, it sends it to all available peers. Both processes have one thing in common: validation. Among the network nodes, there can be both reliable peers and malicious actors that promote fake data. To avoid distortion of the blockchain, each node checks all data received in the network for validity. The validation algorithm sequentially goes through each block and recalculates all the data in search of inconsistencies, such as transactions created by someone other than the wallet owner, distorted block data, etc. Such verification may be slow if there are a large number of blocks. When planning the implementation of a blockchain, it is necessary to carefully choose the means of its creation. It was decided to create all the logic in the Java programming language. To demonstrate the operation and structure of the blockchain, it was decided not to use auxiliary libraries to create it, but to implement it manually. Manual

implementation also provides more flexibility in development and control. The next step is to select auxiliary algorithms and tools for their implementation. The hashing algorithm used was SHA256, as it is quite common and is used in the Bitcoin product, which became the basis for the developed IP. The algorithm can be conveniently implemented using the Java package java.security. RSA was chosen as a cryptographic algorithm primarily because it can be easily implemented using the java.security package. Its main analogue was the ECDSA algorithm, which is based on the properties of elliptic curves, but its implementation in Java is complicated. To generate random numbers, we chose the SecureRandom tool from the java.security package. When choosing the type of network topology, the main criterion was the ability to manage nodes for commercialization. That is why a centralized type was chosen.

Summary of the main material

The first step in using the application is to go to the Products menu. After switching, a blank table of product types and possible options will be displayed. To create a new product type, click the ‘Units’ button. This will open a modal window where you can enter data for the new product. A special feature of this window is the ability to manually enter product type keys or automatically generate them using the ‘Generate Keys’ button. As a result of creating several product types, the window for viewing these types will look as shown in Fig. 2 below. At this stage, you need to select a specific type of product and press the "Units" button to go to specific units of goods (Fig. 3). Each physical item is a product unit. Although different units may have the same product type, each product unit has unique identifiers to distinguish the required product. Create several product units using the "Add" button (Fig. 4). Both the product type creation window and the product unit creation window have a defined key for automatic field generation. In this case, the "Public Info" and "Private Info" fields must be saved, they will be needed later (Fig. 5).

This is what the product units scene looks like after creating several instances. The next step is to add the created data to the blockchain. To do this, return to the main menu and open the transaction scene using the ‘Transactions’ button. The only thing visible in this scene will be an empty transaction table, as no transactions have been created yet. To automatically create the necessary transactions, click the ‘Refresh’ button (Fig. 6).

In Fig. 6, there are four automatically generated transactions. Three of them correspond to the three created units of goods and have a balance of 1, and the last

Id	Public Key	Private Key	Description
1	MIGfMA0GCSqGSIb3DQEBAQUAA4GNADCBiQKBgQ...	MIICdwIBADANBgkqhkiG9w0BAQEFAASCAmEwggJd...	{"name": "Xiaomi Redmi Note 7 Pro","specs":"%3Ch3%3ENETWORK%3C...
2	MIGfMA0GCSqGSIb3DQEBAQUAA4GNADCBiQKBgQ...	MIICdgIBADANBgkqhkiG9w0BAQEFAASCAmAwggJc...	{"name": "Apple Iphone 7","specs":"%3Ch3%3ENETWORK%3C%2Fh3%3E...
3	MIGfMA0GCSqGSIb3DQEBAQUAA4GNADCBiQKBgQ...	MIICdQIBADANBgkqhkiG9w0BAQEFAASCAl8wggJbA...	{"name": "Apple Iphone 8","specs":"%3Ch3%3ENETWORK%3C%2Fh3%3E...
4	MIGfMA0GCSqGSIb3DQEBAQUAA4GNADCBiQKBgQ...	MIICdwIBADANBgkqhkiG9w0BAQEFAASCAmEwggJd...	{"name": "Apple Iphone X","specs":"%3Ch3%3ENETWORK%3C%2Fh3%3...
5	MIGfMA0GCSqGSIb3DQEBAQUAA4GNADCBiQKBgQ...	MIICdgIBADANBgkqhkiG9w0BAQEFAASCAmAwggJc...	{"name": "Samsung Galaxy S21","specs":"%3Ch3%3ENETWORK%3C%2Fh...

Fig. 2. Scene of product types.

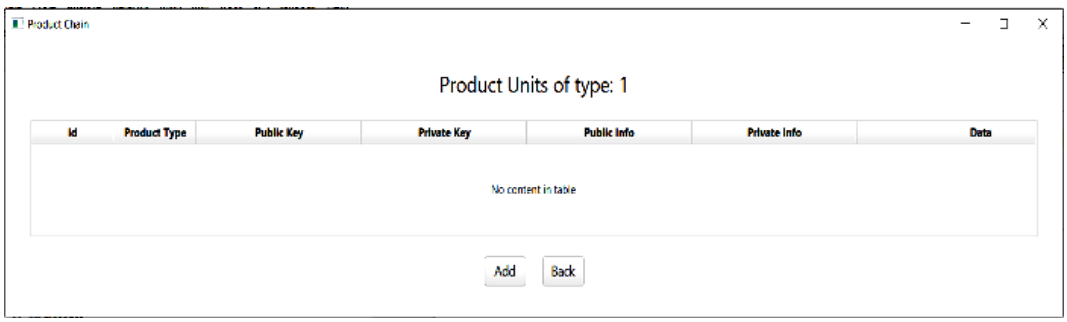


Fig. 3. Product unit scene.

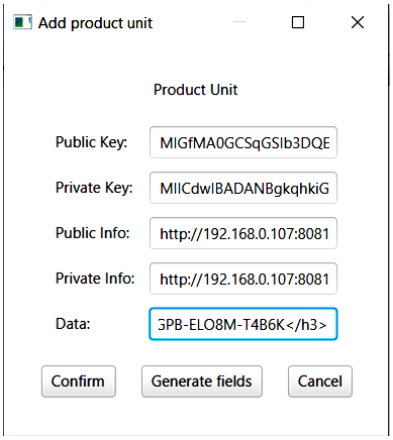


Fig. 4. Product unit creation window.

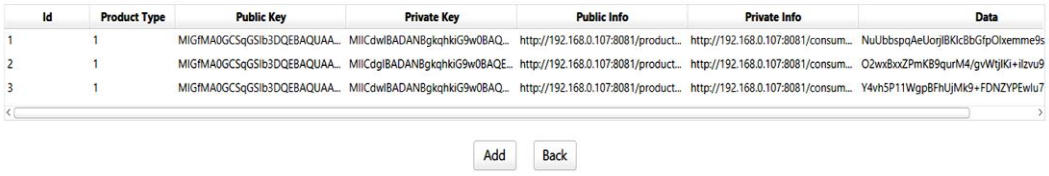


Fig. 5. Product unit scene.

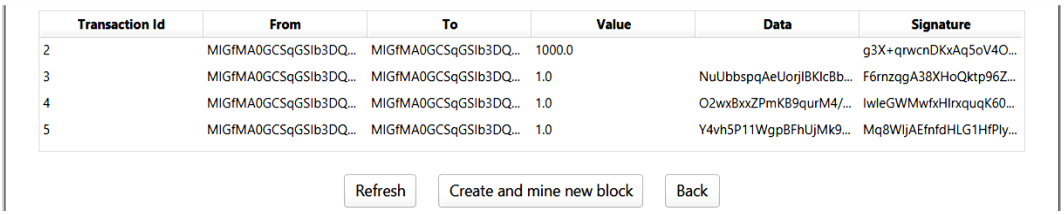


Fig. 6. Transaction viewing scene.

transaction, with a balance of 1000, corresponds to the transfer of funds from the base wallet to a specific type of goods. This is necessary so that units can be created from a type of goods. After such a transfer, it is possible to create 1,000 units of one type of goods. To create a block based on the transaction data and add it to the blockchain, click the ‘Create and mine new block’ button, which will open the ‘Create Block’ modal window (Fig. 7). The creation of a block is visually divided into three stages:

- Notification of how many transactions will be added to the created block. This value depends on the number of transactions created and the maximum possible number of transactions in a block. In this example, the maximum number of transactions is set to 4, so all transactions are included in the block. If there are more transactions, it is necessary to repeat the block creation until all transactions are applied.
- The mining stage is the longest stage, depending on the set complexity.
- Block creation status can be either successful or unsuccessful.

After creating a block, one can proceed to the main menu and the blockchain stage by pressing the ‘Blocks’ button (Fig. 8). On this stage, it is possible to view all blocks and values recorded in the blockchain. By default, this list is organized into a tree structure, sorted by blocks in descending order, and all its values collapsed. Use the ‘Refresh’ button to update the list of blocks to the current one. Now the manufacturer needs to generate QR codes using any generator based on the saved ‘Public info’ and ‘Private Info’ fields and

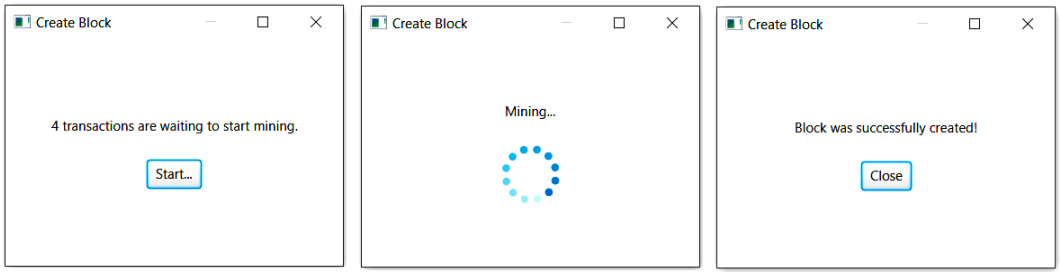


Fig. 7. Stages of creating a block.

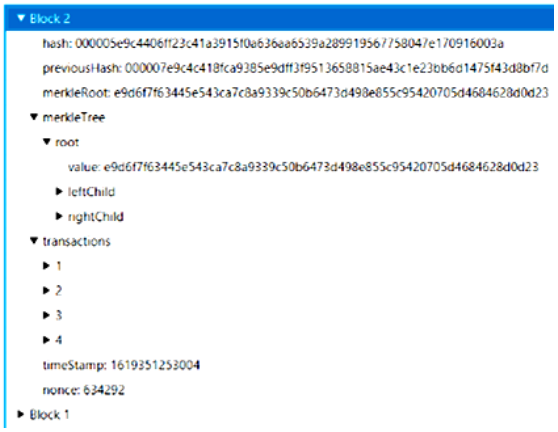


Fig. 8. Blocks viewing scene.

place them on the product packaging (Fig. 9). The control example continues from the buyer's side, and the first step is to scan the public code. Upon receiving the product, before unpacking it, the buyer can scan the external QR code and find out whether the product has already been used and whether the code leads to the correct manufacturer.

Fig. 10a shows what this page looks like on a mobile device. From the information obtained, it is possible to determine the correct manufacturer's website, verify that the characteristics of the actual product match the information on the website, and confirm the status of the product as unused. Therefore, it can be concluded that the product is authentic. Otherwise, the product is determined to be counterfeit, and the buyer can either refuse to purchase the product or contact the appropriate consumer protection authorities. Since this product is authentic, the buyer opens the package and scans the internal QR code. As a result of scanning the internal QR code, the buyer receives information about the specific product, such as activation codes, promo codes, etc. Here comes the second

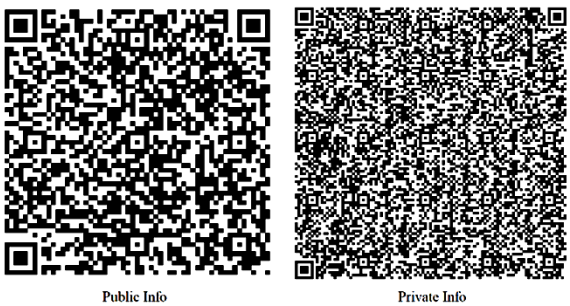


Fig. 9. Generated QR codes.

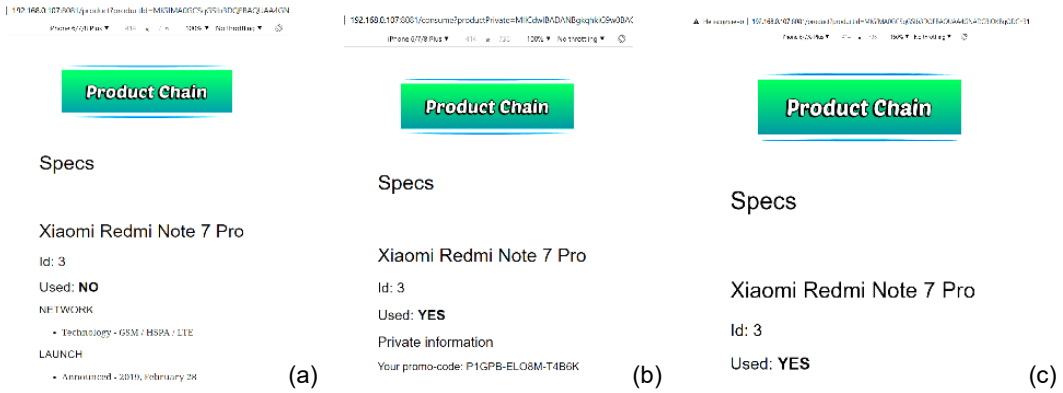


Fig. 10. Transition by a) Public Info, б) Private Info та c) re-transition by Private Info.

authentication factor: the buyer checks whether the product ID from the public and private QR codes match, and then scans the public QR code again, which changes the status of the product to 'used.' After scanning the external QR code again, the buyer sees that the product is now used. At this point, the authenticity check is complete, and the buyer can start using the product. The control example continues from the manufacturer's side. When the private link is clicked, a consumption transaction is created in the manufacturer's Product Chain app, as shown in Fig. 11, which must also be added to the block and then to the blockchain.

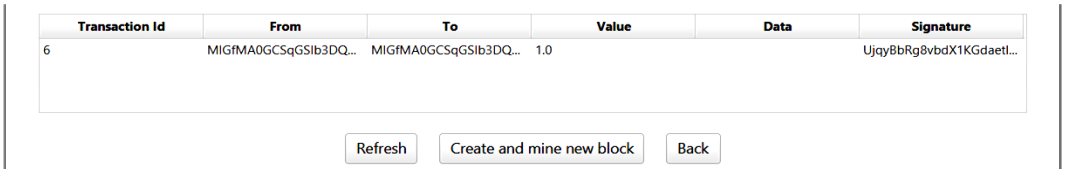


Fig. 11. Transaction of consumption of goods.

Since the application is connected to a P2P server, it can send and receive data from other nodes. Let's start another process of the manufacturer's client application running on other ports using the command: *java -DSERVER_PORT=8001 -DCLIENT_PORT=8102 -DWEB_SERVER_PORT=8082 -jar ProductChain.jar.* The new application will no longer have an empty blockchain but will download the existing blockchain from the first node (Fig. 12).

Fig. 12. Two processes of the Product Chain application running simultaneously/

Both processes contain the same number of blocks. From this point on, after each new block is mined, this block is distributed throughout the blockchain network so that all its nodes have the most up-to-date state. However, the same feature is also a difficulty because if a node receives a new block during the mining process, it will have to cancel mining and start creating the block again. To demonstrate this phenomenon, we start mining simultaneously on two processes (Fig. 13). According to the conclusions of each block, the mining of the first process was unsuccessful, while the second was successful. In other words, the first process received a new block during mining and stopped it in order to add the block to the local copy of the blockchain and recreate the transactions.

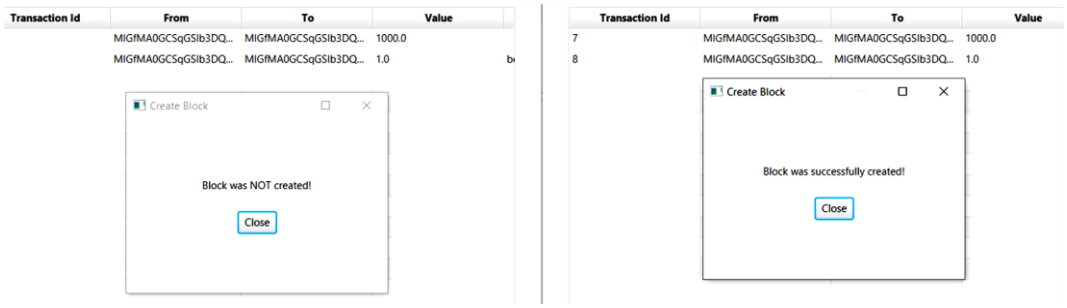


Fig. 13. The result of the simultaneous mining of blocks of two processes

To optimize further work with the program, as well as to search for opportunities to increase performance by changing the working environment, the performance aspects of the system are studied depending on the adjustment of certain parameters. The first such parameter is mining complexity. This parameter determines the number of zeros that must be at the beginning of the hash in order for the work to be considered confirmed. Thus, the complexity range from 1 to 7 is investigated (at complexity 8 and above, mining on the working device takes too long). The results of the analysis are shown in Fig. 14.

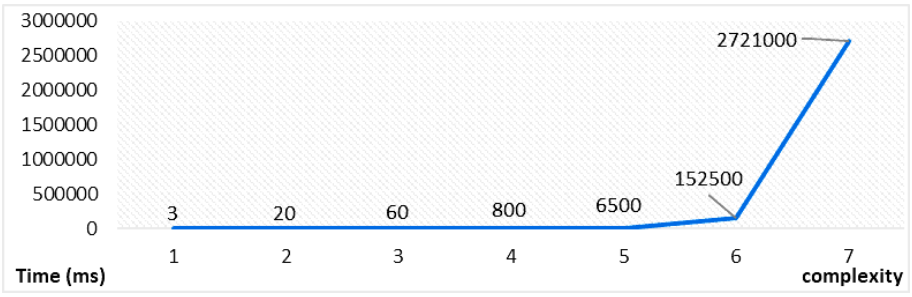


Fig. 14. Graph of the dependence of mining time on its complexity.

The time is shown in milliseconds. This graph shows that each subsequent complexity increases the total time several times. This is due to a significant increase in the number of operations that need to be performed during mining. Of course, mining also depends on the power of the technical equipment. Real devices that mine on a regular basis are a bunch of fancy processors and graphics cards. The dependence of mining time on processor frequency is shown in Figure 15. The analysis was performed for mining complexity - 5. The data shows that additional processor power can significantly reduce mining time. The next subject of research is block creation. The main resource load at this stage is the calculation of hash functions for transactions, for the block, and the construction of a hash tree. The results of the study are shown in Figure 16. From the data obtained, it

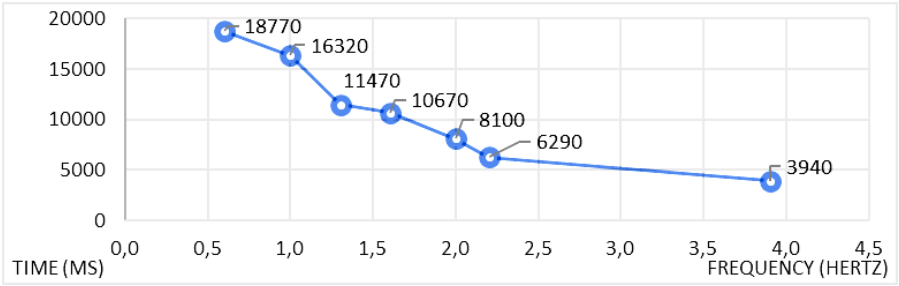


Fig. 15. Graph of mining time versus processor frequency.

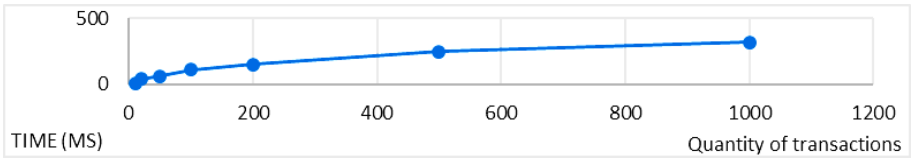


Fig. 16. Graph of block creation time vs. number of transactions.

can be concluded that the number of transactions in a block is its main resource load. This should be taken into account when selecting the maximum number of transactions per block.

The last study was to determine the dependence of blockchain validation time on the number of blocks. The results of this study are shown in Figure 17. The graph shows a linear dependence of the blockchain validation time on the number of blocks. This indicates that the validation process is not overloaded with resource-dependent operations. As the blockchain expands over time, you can expect the expected duration of blockchain validation to be linearly dependent on.

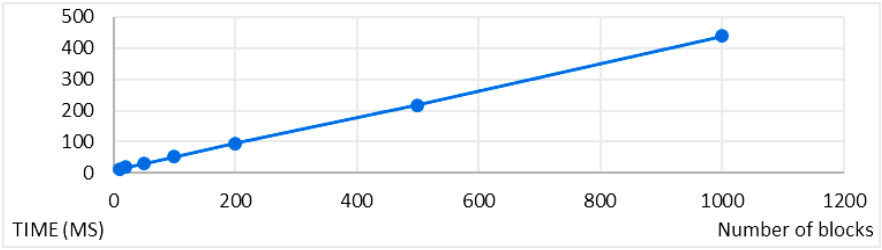


Fig. 17. Graph of the dependence of blockchain validation time on the number of blocks.

RESULTS AND DISCUSSION

As a result of the work, an information system for verifying the authenticity of goods based on blockchain technology was developed. All necessary modules for a complete demonstration of the system were developed, including blockchain modules, a manufacturer's user interface, and a web server. A P2P server was also implemented, allowing the developed software applications to interact. The issues were studied, a system analysis was carried out, and technical means of implementation were selected. After that, the system was developed and its behavior was studied. The work also included a study of the performance aspects of the developed system depending on the adjustment of four parameters: mining complexity, mining time dependence on technical capabilities, block creation time dependence on the number of transactions, and blockchain validation time dependence on the number of blocks. The mining complexity was analyzed on a scale from 1 to 7 (at complexity 8 and above, mining on a working device takes too long). This parameter determines the number of zeros that must be at the beginning of the hash in order for the work to be considered confirmed. Each subsequent complexity increases the total time several times. This is due to a significant increase in the number of operations that must be performed during mining. Of course, mining also depends on the power of the technical equipment. Real devices that mine on a regular basis are a complex of advanced processors and video cards. An analysis of the dependence of mining time on processor frequency was performed for mining difficulty - 5.

CONCLUSION

Additional processor power can significantly reduce mining time. The next study is an analysis of the dependence of block creation time on the number of transactions. The main

resource load at this stage is the calculation of hash functions for transactions, for a block, as well as the construction of a hash tree. From the data obtained, it can be concluded that the number of transactions in a block is its main resource load. This should be taken into account when selecting the maximum number of transactions per block. The last study is to identify the dependence of blockchain validation time on the number of blocks. The data obtained shows a linear dependence of blockchain validation time on the number of blocks. This indicates that the validation process is not overloaded with resource-dependent operations. When expanding the blockchain over time, the expected duration of blockchain validation can be calculated according to the linear dependence.

ACKNOWLEDGMENTS AND FUNDING SOURCES

This publication was prepared with the grant support of the National Research Foundation of Ukraine, project registration number 187/0012 dated 08.01.2024 (2023.04/0012) "Development of an information system for automatic detection of sources of disinformation and unreliable behavior of chat users" within the framework of the competition "Science for Strengthening Ukraine's Defense Capability".

REFERENCES

[1] Li X., Wang C. The technology and economic determinants of cryptocurrency exchange rates: The case of Bitcoin. Decision support systems, 2017, 95: 49-60. https://doi.org/10.1016/j.dss.2016.12.001

[2] Чубенко А.Г., Лошицький М.В., Павлов Д.М., Бичкова С.С., Юнін О.С. Термінологічний словник з питань запобігання та протидії легалізації доходів, одержаних злочинним шляхом, фінансуванню тероризму, фінансуванню розповсюдження зброї масового знищення та корупції, К.: Ваіте, 2018. – 826 с.

[3] Allison. Behind the industry of counterfeit products in China and lawsuit success cases, 2021. URL: https://daxueconsulting.com/counterfeit-products-in-china/

[4] Butticè, V., Caviggioli, F., Franzoni, C., Scellato, G., Stryszowski, P., & Thumm, N. Counterfeiting in digital technologies: An empirical analysis of the economic performance and innovative activities of affected companies. Research Policy, 49(5) (2020) 103959. https://doi.org/10.1016/j.respol.2020.103959

[5] Richter F. The Industries Most Affected by Counterfeit Products, 2019. URL: https://www.statista.com/chart/17410/counterfeit-and-pirated-products-by-category/.

[6] Prokipchuk, O., Chyrun, L., Bublyk, M., Panasyuk, V., Yakimtsov, V., & Kovalchuk, R. (2021). Intelligent System for Checking the Authenticity of Goods Based on Blockchain Technology. In MoMLeT+ D, pp. 618-665.

[7] Kumar R., Tripathi R. Traceability of counterfeit medicine supply chain through Blockchain. In: 11th international conference on communication systems & networks (COMSNETS). IEEE, 2019. p. 568-570. https://doi.org/10.1109/COMSNETS.2019.8711418

[8] Alipour S. Ninety Eight Per Cent Of Fake Or Lookalike IPhone Chargers Put Consumers At Risk Of Lethal Electric Shock And Fire, 2017. URL: https://www.electricalsafetyfirst.org.uk/media-centre/press-releases/2017/12/ninety-eight-per-cent-of-fake-or-lookalike-iphone-chargers-put-consumers-at-risk-of-lethal-electric-shock-and-fire/.

[9] Casassus B. Health agency reveals scourge of fake drugs in developing world, 2017. URL: https://www.nature.com/news/health-agency-reveals-scourge-of-fake-drugs-in-developing-world-1.23051.

[10] Hamelin N., Nwankwo S., El Hadouchi R. Faking brands': consumer responses to counterfeiting. Journal of Consumer Behaviour, 12(3), (2013) 159-170. https://doi.org/10.1002/cb.1406

[11] Zhang L. Platformizing family production: The contradictions of rural digital labor in China. The Economic and Labour Relations Review, 32(3), (2021). 341-359. https://doi.org/10.1177/10353046211037093
[12] Horstmann C. S., Cornell G. Core Java 2: Fundamentals (Vol. 1). Prentice Hall Professional, 2001.
[13] Martin R. Clean Architecture: A Craftsman's Guide To Software Structure And Design. U.S.A.: Pearson, 2017.
[14] Freeman E., Robson E., Bates B., Sierra K. Head First Design Patterns: A Brain-Friendly Guide. " O'Reilly Media, Inc.", 2004.
[15] Walls C. Spring in Action. New York: Manning Publications, 2018.
[16] Walls C. Spring Boot in Action. New York: Manning Publications, 2018.
[17] Grinev S. Mastering JavaFX 10: Build advanced and visually stunning Java applications. New York: Packt Publishing, 2018.
[18] Chaitanya S., JSON Tutorial, 2015. URL: https://beginnersbook.com/2015/04/json-tutorial/.
[19] Aravind M. Gson Library, 2017. URL: https://medium.com/@manuaravindpta/gson-library-b7d4ef0381e2.
[20] Maven: The Definitive Guide: The Definitive Guide – New York: O'Reilly Media, 2008.
[21] Schlager R. The OSI Model: simply explained. New York: CreateSpace Independent Publishing Platform, 2013.
[22] Gourley D., Totty B., Sayer M., Aggarwal A. HTTP: The Definitive Guide. New York: O'Reilly Media, 2002.
[23] Kurniawan B. How Tomcat Works. New York: Brainy Software, 2005.
[24] Aquino C., Gandee T. Front-End Web Development. New York: Big Nerd Ranch Guides, 2016.
[25] Kommadi B. IntelliJ vs Eclipse Complete IDE Comparison, 2019. URL: https://medium.com/@bhagvankommadi/hi-team-6d2dee22d8b2.
[26] Reed J. Blockchain: The Essential Guide to Understanding the Blockchain Revolution. New York: CreateSpace Independent Publishing Platform, 2016.
[27] Walker G. Block Header. A summary of the data in the block, 2016. URL: https://learnmeabitcoin.com/technical/block-header.
[28] Walker G. Merkle Root. A fingerprint for all the transactions in a block, 2016. – URL: https://learnmeabitcoin.com/technical/merkle-root.
[29] Canty R. Understanding Cryptography with RSA, 2020. URL: https://jryancanty.medium.com/understanding-cryptography-with-rsa-74721350331f.
[30] Nagpal A. How to create your own decentralized file sharing service using python, 2018. URL: https://medium.com/@amannagpal4/how-to-create-your-own-decentralized-file-sharing-service-using-python-2e00005bdc4a.

COMPLIANCE WITH ETHICAL STANDARDS

Conflict of Interest: The authors declare that the research was conducted in the absence of any conflicts of interest.

AUTHOR CONTRIBUTIONS

Conceptualization, [V.V.]; methodology, [V.V., O.P.]; validation, [M.N.]; formal analysis, [V.V.].; investigation, [O.P.]; resources, [O.P.]; data curation, [M.N.]; writing – original draft preparation, [M.N.]; writing – review and editing, [M.N.]; visualization, [O.P.] supervision, [R.R.]; project administration, [R.R.]; funding acquisition, [R.R.].

All authors have read and agreed to the published version of the manuscript.

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ДЛЯ АУТЕНТИФІКАЦІЇ ДАНИХ НА ОСНОВІ БЛОКЧЕЙНУ

***Вікторія Висоцька*[1], *Олег Прокіпчук*[1], *Марія Назаркевич*[2]*, *Роман Романчук*[1]**

[1]Кафедра інформаційних систем та мереж
Національного університету «Львівська політехніка»,
вул. Степана Бандери 12, м. Львів, 79013, Україна
[2]Кафедра радіофізики та комп'ютерних технологій
Львівського національного університету імені Івана Франка,
вул. Драгоманова 50, Львів, 79005, Україна

АНОТАЦІЯ

Вступ. В статті описано процес розроблення технології інформаційної підтримки перевірки автентичності товарів на основі блокчейн. В роботі здійснено дослідження швидкодії розробленої системи залежно від корегування чотирьох параметрів: складність майнінгу, залежність часу майнінгу від потужностей технічного забезпечення, залежність часу створення блоку від кількості транзакцій та залежність часу валідації блокчейну від кількості блоків.

Матеріали та методи. Аналіз складності майнінгу проводився в межах від 1 до 7 (на складності 8 та більше, майнінг на робочому пристрої відбувається занадто довго). Цей параметр визначає кількість нулів, що мають бути на початку хешу, щоб вважати роботу підтвердженою. Далі складність збільшує загальний час роботи у декілька разів. Це пов'язано із збільшенням кількості операцій, що необхідно виконати при майнінгу. Звісно майнінг також залежить від потужностей технічного забезпечення, які мають бути комплексом передових процесорів та відеокарт. Аналіз залежності часу майнінгу від частоти процесора проведений для складності майнінгу - 5.

Результати. Додаткова потужність процесора може значно скоротити час майнінгу. Наступним дослідженням є аналіз залежності часу створення блоку від кількості транзакцій. Основне ресурсне навантаження цього етапу є обчислення хеш-функцій для транзакцій для блоку, а також побудова хеш-дерева. З отриманих даних можна зробити висновок, що кількість транзакцій блоку є основним його ресурсним навантаженням. Це варто враховувати при виборі максимальної кількості транзакцій до блоку.

Висновки. Останнім дослідженням є виявлення залежності часу валідації блокчейну від кількості блоків. З отриманих даних випливає лінійна залежність часу валідації блокчейну від кількості блоків. Це сигналізує про те, що процес валідації не перевантажений ресурсозалежними операціями. При розширенні блокчейну з часом можна розраховувати на очікувану тривалість валідації блокчейну за лінійною залежністю.

Ключові слова – мережевий зв'язок, інформаційна система, контрафактний товар, технологія блокчейн, життєвий цикл продукту, інтелектуальна система пошуку інформації.

Received / Одержано	Revised / Доопрацьовано	Accepted / Прийнято	Published / Опубліковано
24 February, 2025	18 March, 2025	24 March, 2025	26 June, 2025

Electronics and Information Technologies, 2025, 30, 75–88
http://publications.lnu.edu.ua/collections/index.php/electronics/index

UDC: 004.896 + 004.415

BENCHMARKING GAUSS AND LAPLACE ARTIFICIAL POTENTIAL FIELD APPROACHES FOR REAL-TIME OBSTACLE AVOIDANCE IN VIRTUAL SCENARIOS

Ihor Berizka*, Ivan Karbovnyk
Department of Radiophysics and Computer Technologies,
Ivan Franko National University of Lviv,
107 Tarnavskoho St., Lviv 79013, Ukraine

Ihor Berizka, Ivan Karbovnyk. (2025). Benchmarking Gauss and Laplace Artificial Potential Field Approaches for Real-Time Obstacle Avoidance in Virtual Scenarios. *Electronics and Information Technologies*, 30, 75–88. https://doi.org/10.30970/eli.30.6

ABSTRACT

Background. Autonomous mobile robots require robust real-time obstacle avoidance algorithms to navigate dynamic environments safely. The Artificial Potential Field (APF) method remains widely adopted for local path planning due to its computational efficiency and conceptual simplicity. However, conventional implementations suffer from two well-documented limitations: local minima and computational inefficiencies. This study investigates two probabilistic APF variants – a Gaussian formulation (ODG-PF) and a Laplace-based approach to address these limitations

Materials and Methods. A comparative framework was developed using ROS2/Gazebo with TurtleBot3 as target platform. The Gaussian APFM (ODG-PF) and Laplace APFM were mathematically modeled, with key differences in their repulsive force calculations: Gaussian uses squared terms, while Laplace employs absolute values. Both methods were tested in identical static environments with 25 repeated runs (28 steps each). Performance metrics included computational time and path length, analyzed via boxplots, kernel density estimation, and Mann-Whitney U tests ($p<0.05$).

Results and Discussion. The Laplace APFM demonstrated superior efficiency, with 34% faster median execution time (68 μs vs. 104 μs) and tighter interquartile range (28 μs vs. 52 μs). Its unimodal time distribution contrasted with the Gaussian's bimodal pattern, attributed to simpler arithmetic operations. While both methods achieved collision-free navigation, Laplace generated statistically shorter paths ($p=0.0001$), though with marginally higher variability. The Gaussian method's squaring operations introduced computational overhead without navigational benefits.

Conclusion. The Laplace-based APFM outperforms its Gaussian counterpart in computational speed and path optimization, making it ideal for resource-constrained systems. These findings suggest that simpler mathematical formulations can yield superior real-world performance in obstacle avoidance applications. Future work should validate these findings in dynamic environments and explore hybrid implementations with global planners.

Keywords: cyber-physical system, information technologies, obstacle avoidance, mobile robotic platforms, IoT concepts, wheeled mobile platform

INTRODUCTION

In the field of robotics, a significant area of research and development is focused on autonomous mobile robots. These advanced systems are designed to navigate

independently and make context-specific decisions in real time. This capability enables them to operate without human intervention, responding adaptively to dynamic environments based on continuous sensory input.

Autonomous mobile robots are employed across a wide range of applications. Examples include service robots in hospitality environments, such as waiter robots that deliver food and beverages, and transport robots used in industrial settings to move goods efficiently. A particularly prominent example is autonomous vehicles, commonly referred to as self-driving cars. These systems integrate sophisticated sensor arrays and computational algorithms to interpret their surroundings, allowing them to navigate complex traffic scenarios and reach destinations without human control.

These examples underscore the wide-ranging applicability of autonomous mobile robots and emphasize their potential to transform numerous sectors, including hospitality, logistics, and the automotive industry. The continued development and optimization of these systems remain a central focus within the robotics research community, with ongoing efforts directed toward advancing their functional capabilities and broadening their domains of deployment.

One of core software components in autonomous mobile robots is the set of algorithms dedicated to path planning and obstacle avoidance. These algorithms are essential for enabling functionalities such as autonomous parking, evasive maneuvers in emergency scenarios, and ultimately, full autonomy in navigation and control.

Path planning is typically categorized into two main types: global and local. Global path planning relies on data from Geographic Information Systems (GIS) in conjunction with global localization techniques. This approach requires the robot to possess a comprehensive, large-scale understanding of its environment, enabling navigation over extended distances - such as traversing urban areas or intercity routes.

In contrast, local path planning requires only the robot's relative position and real-time perception of obstacles within its immediate environment. This form of planning focuses on short-range navigation and dynamic interactions with the surroundings, such as avoiding pedestrians on a sidewalk or maneuvering around other vehicles in traffic. Local path planning is crucial for ensuring safe and responsive behavior in unpredictable and rapidly changing environments.

A wide range of algorithms has been developed to address both global and local path planning challenges. Each of these algorithms presents distinct advantages and limitations, and their selection can have a substantial influence on the overall efficiency, reliability, and safety of autonomous navigation. Comprehensive reviews of these path planning techniques, including their underlying methodologies and application domains, can be found in the literature [1, 2, 3, 4].

Obstacle detection and avoidance constitute a critical component of local path planning algorithms, serving a vital role in ensuring the safety of both the autonomous system and its surrounding environment. This area has been the focus of extensive research over several decades, leading to the development of numerous methodological approaches. Many of these techniques have demonstrated practical effectiveness and have been successfully implemented in real-world applications.

To effectively avoid collisions, a robot must not only detect obstacles but also dynamically recalculate its path and modify its trajectory in real time. Real-time responsiveness is essential for navigating complex and dynamic environments safely and efficiently.

The initial step in obstacle avoidance involves the robot detecting potential obstacles using its onboard sensors. Once an obstacle is identified, the system must generate a new trajectory that enables safe navigation around the object. This alternative path must be computed with minimal latency to ensure that the robot can adjust its motion in real time, thereby preventing collisions and maintaining smooth path.

Additionally, the robot must be able to adapt to dynamic changes in its environment. For example, if a new obstacle unexpectedly occurs in its path, the robot needs to rapidly

detect the obstacle, compute an alternative route, and adjust its trajectory in real time to ensure safe navigation.

This demonstrates the complexity inherent in autonomous mobile robots and highlights the critical importance of ongoing research in this domain. The development and optimization of robust obstacle detection and avoidance algorithms remain central to robotics research, with the objective of improving the safety, reliability, and efficiency in terms of path length, computational time and resources of autonomous systems.

The obstacle avoidance problem can be framed as follows: A robot is situated within an unknown environment. The robot is expected to reach a specified target location or following a goal direction, all while navigating around any obstacles that may occur on its path. More detailed overview of the problem is presented in [1, 2].

MATERIALS AND METHODS

The Artificial Potential Field Method (APFM) is a well-established approach in robotics, widely applied in path planning and obstacle avoidance tasks. Originally proposed by Khatib [5], the method models the robot's environment as a virtual potential field, where the target location produces an attractive force, while obstacles produce repulsive forces. The robot, treated as a particle under the influence of these virtual forces, is guided toward the goal while being repelled from surrounding obstacles. Its motion is determined by the resultant force vector calculated as the superposition of these attractive and repulsive components. The mathematical formulation of these forces is represented by Eqs. (1)-(3).

$$\mathbf{f}_{\text{total}} = \mathbf{f}_{\text{rep}} + \mathbf{f}_{\text{attr}}, \tag{1}$$

$$\mathbf{f}_{\text{attr}} = k_{\text{attr}} \cdot \frac{\mathbf{r}_{\text{goal}} - \mathbf{r}}{\left|\mathbf{r}_{\text{goal}} - \mathbf{r}\right|}, \tag{2}$$

$$\mathbf{f}_{\text{rep}} = \begin{cases} -k_{\text{rep}} \cdot \sum_{i=1}^{n} \left(\frac{1}{d_i} - \frac{1}{d_{\max}}\right) \cdot \mathbf{s}_i, & if\ d_i < d_{\max} \\ 0, & if\ d_i \geq d_{\max} \end{cases}, \tag{3}$$

where $\mathbf{s}_i = (\mathbf{r} - \mathbf{o}_i)/|\mathbf{r} - \mathbf{o}_i|$, $\mathbf{r}_{\text{goal}}$ is the position vector of the goal point and $\mathbf{r}$ is the position vector of the vehicle, $\mathbf{o}_i$ – the position vector of each obstacle.

Despite its widespread use, the traditional Artificial Potential Field method (APFM) has several limitations. One of the primary issues is the occurrence of local minima, where the robot may become trapped in a position that is not the target, as the attractive force towards the goal and the repulsive forces from obstacles cancel each other out. Additionally, the method can result in situations where the target becomes unreachable, or the robot follows inefficient paths due to suboptimal force interactions [2, 6]. To overcome these issues, improved versions of the APF method have been proposed. One of such modifications of classic APFM is based on probabilities. In the following subsections, we examine these modifications within the mathematical models of the method in detail.

Mathematical model of Gauss APFM

The Obstacle-Dependent Gaussian Potential Field (ODG-PF) method was developed and implemented to facilitate obstacle detection and assess the probability of collision and is presented in detail in [7]. This study introduces a novel approach to calculating attractive and repulsive fields, as well as an innovative direction decision strategy. Comprehensive simulations and experimental evaluations were conducted, comparing the ODG-PF method with other potential field-based obstacle avoidance techniques. The results demonstrate that the ODG-PF method outperforms existing approaches in the majority of tested scenarios.

During the development of the mathematical model, the authors introduce a pair of values (θ, d), where θ represents the measurement angle in degrees, and d denotes the distance to an object in meters. To obtain data in this format, the authors propose using a movable ultrasonic distance sensor. Alternatively, a one-dimensional LiDAR can be employed to scan the working environment.

At the initial stage, it is necessary to define a threshold distance d_{thr}. All objects located closer to the robot than d_{thr} are considered as obstacles. This parameter can significantly influence the performance of the algorithm and therefore requires empirical tuning. Obstacles are defined using additional parameters represented as a pair $(\theta_{\text{start}}, \theta_{\text{end}})$, where θ_{start} is the angle at which the obstacle is first detected, and θ_{end} is the angle at which the detection of the same obstacle ends. This approach enables obstacle identification from one-dimensional LiDAR input data. In addition to angular boundaries, it is necessary to compute supplementary parameters for each obstacle: d_k the average distance to the k-*th* obstacle; $\Phi_k = \theta_{\text{end}} - \theta_{\text{start}}$ is the angular width occupied by the obstacle. Additionally, the robot is modeled as a square with side length w_{robot} in meters.

Although in many experiments or computer simulations the dimensions of the robot are often neglected - treating it as a point mass to simplify simulations and mathematical models - considering the robot's physical dimensions is essential for improving the model's practicality and aligning it more closely with real-world conditions. While this increases the complexity of the model to some extent, it also enhances its practical value. The authors propose a hybrid approach: the robot's dimensions are taken into account, but they are incorporated into the model by modifying the perceived size of surrounding obstacles using specific formulas. In this way, obstacles are effectively enlarged from the robot's perspective, while the robot itself continues to be modeled as a point mass.

To account for the dimensions of the robot, it is necessary to recalculate the angle according to the following equation:

$$\varphi_k = 2\sigma_k = 2\text{atan}\left[d_k \cdot \tan\left(\frac{\Phi_k}{2}\right) + \frac{w_{\text{robot}}}{2}, d_k\right] \tag{4}$$

The next step involves calculating the repulsive force exerted by each obstacle according to Eq. (5).

$$f_k(\theta_i) = A_k \cdot \exp\left[-\frac{(\theta_k - \theta_i)^2}{2\sigma_k^2}\right], \tag{5}$$

where θ_k denotes the central angle of the obstacle, and σ_k represents half of the angular width occupied by the obstacle.

The coefficient A_k is selected such that the Gaussian function fully encompasses the obstacle and is computed according to Eq. (6).

$$A_k = \tilde{d}_k \cdot \exp(0.5), \tag{6}$$

where $\tilde{d}_k = d_{\max} - d_k$, and $d_{\max}$ is sensor range distance.

As with other artificial potential field-based methods, the repulsive force f_k represents the field generated by the k-*th* obstacle. Thus, the overall repulsive field is computed as the sum of the repulsive forces from all individual obstacles according to the Eq. (7).

$$f_{\text{rep}}(\theta_i) = \sum_{k=1}^{n} A_k \cdot \exp\left[-\frac{(\theta_k - \theta_i)^2}{2\sigma_k^2}\right] \tag{7}$$

The next step is the computation of the attractive field using Eq. (8). This field represents the force that draws the robot toward the specified direction θ_{goal}. Consequently, the resulting field is calculated according to Eq. (9) and determines the safe movement direction for the robot at the current iteration (Eq. (10)). The safe movement direction is computed as the argument at which the total potential field function reaches its minimum. This is a significant modification, as it simplifies the computation compared to the classical approach. Parameter γ is selected experimentally and is set to $\gamma = 0.06$

$$f_{\text{attr}}(\theta_i) = \gamma \cdot |\theta_{\text{goal}} - \theta_i|, \tag{8}$$

$$f_{\text{total}}(\theta_i) = f_{\text{attr}}(\theta_i) + f_{\text{rep}}(\theta_i), \tag{9}$$

$$\theta_{\text{dir}} = \arg\min(f_{\text{total}}). \tag{10}$$

In the traditional method, the movement direction is determined using the arctangent between two vectors, which increases computational complexity - particularly for resource-constrained systems. Another advantage of this approach is its compatibility with data obtained solely from ultrasonic distance sensors or one-dimensional LiDAR sensors. This choice of data format considerably reduces computational load, especially when compared to algorithms that rely on obstacle detection and segmentation from images (i.e., computer vision and image processing techniques). It is also worth noting that the authors conducted experiments in both static and dynamic environments and reported that the algorithm requires no additional adjustments or modifications to operate effectively in the presence of moving obstacles.

Mathematical model of Laplace APFM

Following the analysis of the work presented in [7], we identified opportunities to extend this line of research by exploring alternative probability density functions for modeling the repulsive force. To maintain symmetry in the initial modeling phase, we selected the Laplace distribution due to its simplicity and analytical tractability. In our subsequent work [8], we proposed a mathematical formulation of the repulsive field based on the Laplace function, providing a foundation for further investigation into its applicability and performance in robot navigation tasks.

All preliminary steps described for the artificial potential field method (APFM) based on the Gaussian function remain unchanged. Specifically, the environment is scanned and obstacles are detected, followed by their expansion to account for safety margins. The primary modification lies in the calculation of the repulsive force. By substituting the Gaussian function with the Laplace function, we derived the Eq. (11) to describe the repulsive force within the proposed framework.

$$f_k(\theta_i) = A_k \cdot \exp\left(-\frac{\sqrt{2} \cdot |\theta_k - \theta_i|}{\sigma_k}\right), \tag{11}$$

where θ_k denotes the central angle of the obstacle, and σ_k represents half of the angular width occupied by the obstacle.

Upon analysis of Eq. (11), it becomes evident that this formulation offers computational advantages, as it eliminates the need to square the value inside the exponential term – instead relying solely on the absolute value, which simplifies the calculation.

The coefficient A_k serves the same purpose in shaping the repulsive field as in the Gaussian variation. It is adjusted such that the Laplace function adequately covers the spatial extent of the obstacle, ensuring a precise representation of the repulsive force in its proximity, and is calculated using Eq. (12). The derivation of this coefficient is based on the

so-called sigma rule, which has been adapted for the Laplace probability density function (PDF). This approach allows for the systematic tuning of the function's spread in relation to the dimensions of the obstacle.

$$A_k = \tilde{d}_k \cdot \exp(\sqrt{2}), \tag{12}$$

where $\tilde{d}_k = d_{\max} - d_k$, and $d_{\max}$ is sensor range distance.

To account for the influence of each obstacle, the total obstacle field is determined by summing the repulsive fields produced by all obstacles. Consequently, the resulting function becomes a function of the angle θ_i

$$f_{\text{rep}}(\theta_i) = \sum_{k=1}^{n} A_k \cdot \exp\left(-\frac{\sqrt{2} \cdot |\theta_k - \theta_i|}{\sigma_k}\right). \tag{13}$$

The subsequent stage involves the computation of the attractive field, which follows the same formulation as in the Gaussian-based modification. This field represents the force that draws the robot toward the target direction of movement. When combined with the repulsive field, the attractive field contributes to shaping the robot's trajectory, enabling it to avoid obstacles while steadily progressing toward the goal. The second distinction at this stage lies in the necessity to experimentally adjust the coefficient γ for the Laplace-based formulation. Parameter γ is selected experimentally and is set to $\gamma = 6.36$. Total field produced by the system is calculated using Eq. (9), and the safe direction of robot movement is determined using Eq. (10).

Evaluation framework

In robotics research and development, the use of standardized software frameworks and simulation environments is essential for efficient algorithm design, validation, and deployment. One of the most prominent and widely adopted frameworks is the Robot Operating System (ROS), which provides a comprehensive set of tools, libraries, and conventions for developing modular and scalable robotic applications. ROS enables seamless communication between different software components through a publish-subscribe architecture and supports integration with a wide variety of sensors, actuators, and control algorithms. Its flexibility and strong community support have made it a de facto standard in both academic and industrial robotics.

In parallel, simulation environments play a critical role in the development pipeline by offering safe, controlled, and reproducible testing scenarios. Among these, Gazebo stands out as one of the most robust and feature-rich robotic simulators. It provides realistic physics-based modeling of robot dynamics, environmental interactions, and sensor feedback, making it well-suited for prototyping and validating complex robotic behaviors without the risks or costs associated with physical testing. When used in conjunction with ROS, Gazebo allows for rapid iteration and debugging of robotic systems in a simulated 3D environment, closely mirroring real-world conditions.

These software tools not only streamline the development process but also enable researchers to perform extensive benchmarking and testing under varied environmental conditions, which is crucial for ensuring the reliability and robustness of robotic algorithms, especially those involved in autonomous navigation and obstacle avoidance. A more detailed review of the rationale behind the selection of ROS 2, Gazebo, and the TurtleBot3 platform is provided in [9].

Motivated by these capabilities, in [9] we introduced a dedicated testing framework built upon ROS and Gazebo, specifically designed to evaluate obstacle avoidance strategies. This framework utilizes the TurtleBot3 Burger robot model – a widely used

platform in the ROS ecosystem known for its affordability, modularity, and compatibility with both simulation and real-world deployment. The framework allows for the systematic assessment of obstacle avoidance algorithms in simulated environments that closely approximate the conditions encountered in real-world scenarios. By leveraging this simulation infrastructure, we aim to validate the correctness and performance of our proposed mathematical model prior to its application on physical robotic systems.

The proposed framework creates a virtual indoor environment populated with static obstacles, simulating common navigation challenges typically encountered by mobile robots. Within this simulated environment, the TurtleBot3 Burger robot is located in the center of the room and is tasked with navigating autonomously while avoiding collisions with the surrounding obstacles. Virtual room with obstacles in Gazebo simulator is presented on Fig. 1. The primary objective is to evaluate the robot's ability to traverse space safely and efficiently using the implemented obstacle avoidance algorithms.

In addition to the core simulation setup, the framework includes auxiliary software components for debugging, performance monitoring, and statistical analysis. These tools enable visualization and logging of key data such as LiDAR readings, computed potential fields, and the robot's steering angle at each time step, which are essential for understanding the robot's perception and control behavior. This facilitates the identification of performance issues and the refinement of navigational algorithms. The framework also supports the generation of statistical visualizations, including box plots, kernel density estimates (KDEs), and histograms for specified data. Furthermore, it performs non-parametric Mann–Whitney tests and computes statistical indicators such as the mean, median, standard deviation, min, max, IQR and p-values. The results of this analysis are presented in the following sections.

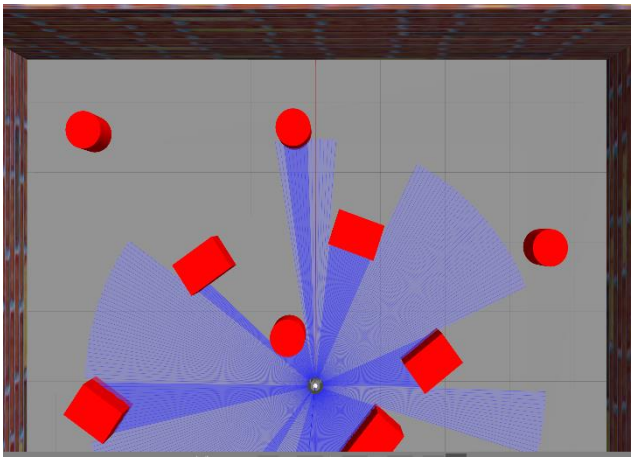


Fig. 1. View of virtual room with obstacles in Gazebo simulator. Robot is located in the center of the room. Obstacles are colored in red.

RESULTS AND DISCUSSION

In our previous work [9], we demonstrated that the Artificial Potential Field Method (APFM) based on the Laplace distribution generates smooth paths and enables reliable obstacle avoidance in environments with static obstacles, achieving collision-free navigation. In the present study, we extend that analysis by performing a comparative statistical evaluation of the Gaussian-based and Laplace-based APFM approaches. Specifically, we compare the total path lengths produced by both algorithms under identical environmental conditions. Additionally, we analyze the computational efficiency of each method.

The angle toward the goal in all the algorithms in this study is fixed and set to $\theta_{\text{goal}} = 0°$. That is to say, there is no goal position but the fixed angle and the purpose of

the obstacle avoidance algorithms in this paper is to do their best to keep the fixed heading without colliding with anything.

The tests were performed on a virtual machine running Ubuntu 20.04, equipped with 8 GB of RAM and 8 virtual CPU cores. The virtual environment was hosted on a physical machine featuring an AMD Ryzen 7 3700X processor, 32 GB of RAM, a Samsung 970 EVO Plus NVMe M.2 SSD, and an NVIDIA RTX 3070 GPU.

Statistical analysis of computational time between Gauss and Laplace APFMs

To assess the performance of two modified versions of the APFM navigation algorithm – Laplace and Gauss – a comprehensive comparative analysis was conducted. Each algorithm was executed 25 times, with each run comprising 28 discrete steps, as shown in Fig. 2, where the highlighted steps correspond to those for which execution time was measured. The primary evaluation metric was the total computational time required per highlighted steps, measured in microseconds. This procedure resulted in a dataset of 700 samples for each method. The statistical outcomes derived from this analysis are presented in Table 1.

In the context of execution time analysis for robotic algorithms, the Interquartile Range (IQR) method for outlier detection and removal was used due to its robustness and distribution-independent nature. Execution time measurements may occasionally include anomalous values caused by transient system states, background processes, or irregular hardware scheduling events - none of which reflect the true performance of the algorithm. The IQR method effectively identifies and excludes these statistical outliers by focusing on the central 50% of the data and filtering out values that lie beyond 1.5 times the interquartile range from the first and third quartiles. This ensures that the resulting statistical analysis reflects the typical behavior of the algorithm, providing more reliable and interpretable performance metrics. Given that our data may not follow a perfectly normal distribution, this non-parametric approach is especially well-suited for maintaining analytical integrity.

Following the outlier filtering step based on the interquartile range (IQR) method, the Gaussian dataset retained 637 samples, while the Laplace dataset retained 654 samples.

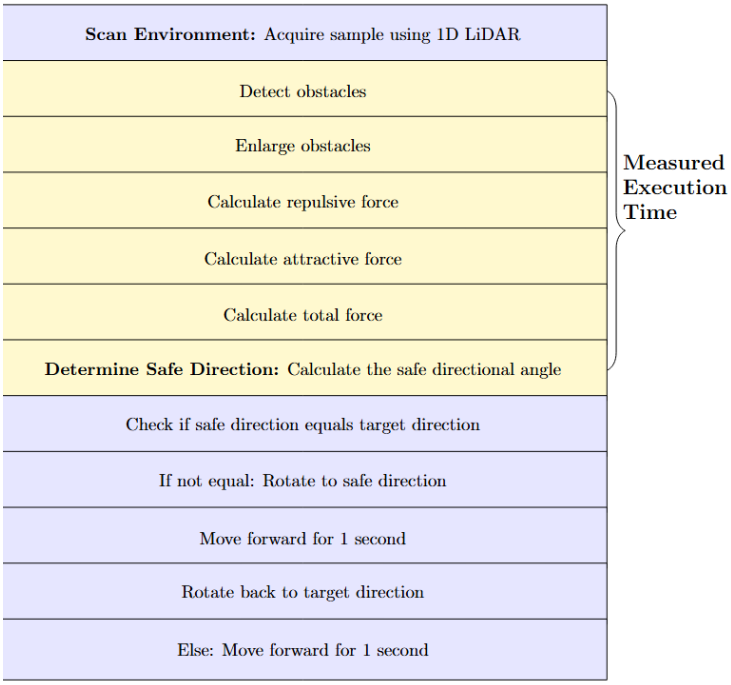


Fig. 2. Pseudocode of general robot navigation algorithm based APFM family algorithms. Steps for which execution time was measured are highlighted.

Table 1. **The calculated statistical values of computational time for Laplace and Gauss APFMs**

APFM variation	Statistical values of computational time, us					
	Mean	Median	Std	Min	Max	IQR
Laplace	76.690	68.000	26.453	26.000	159.000	28.000
Gauss	107.419	104.000	44.909	34.000	237.000	52.000

The relatively small number of excluded data points indicates a high degree of consistency in the measurements. The resulting datasets are therefore more representative of the typical performance of each method, enhancing the reliability and interpretability of the subsequent statistical analysis.

Fig. 3 presents a boxplot comparison of the execution times for the Laplace and Gauss variants of the APFM navigation algorithm. The median execution time for the Laplace algorithm is significantly lower than that of the Gauss variant, indicating superior computational efficiency. Additionally, Laplace exhibits a narrower interquartile range (IQR), suggesting more consistent performance across runs. In contrast, the Gauss algorithm shows greater variability, with a wider IQR and a higher number of outliers, some exceeding 200 microseconds. These results indicate that the Laplace-based implementation not only executes faster on average but also provides more stable timing behavior which might be crucial for real-time applications.

Fig. 4 presents execution time histograms for the Laplace and Gauss variants of the APFM navigation algorithm. The Laplace histogram (left) demonstrates a sharply peaked distribution centered around 65–70 microseconds, with the majority of samples clustered within a relatively narrow range. This indicates both high efficiency and temporal consistency. A small number of outliers are present beyond 100 microseconds, but they remain infrequent.

In contrast, the Gauss histogram (right) displays a broader and more dispersed distribution, with a secondary concentration around 120 microseconds and a notable tail extending beyond 200 microseconds. The wider spread and presence of multiple local peaks suggest greater variability and occasional latency spikes in execution time. These findings align with the results shown in the corresponding boxplot (Fig. 4), confirming the Laplace variant's advantage in both average performance and stability.

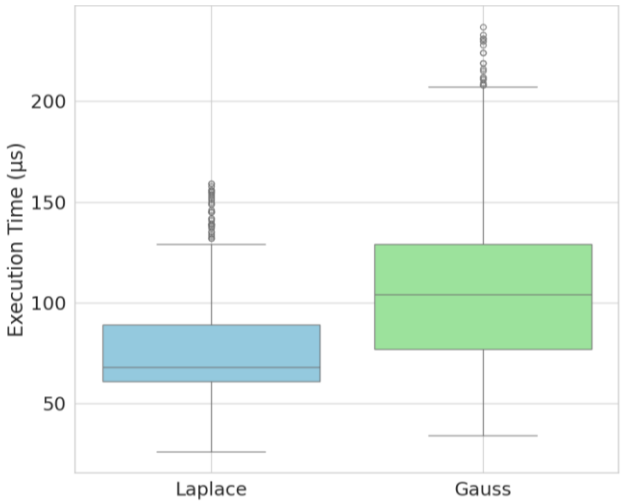


Fig. 3. Boxplot visualization of the experiment results for computational time of Laplace and Gauss modifications.

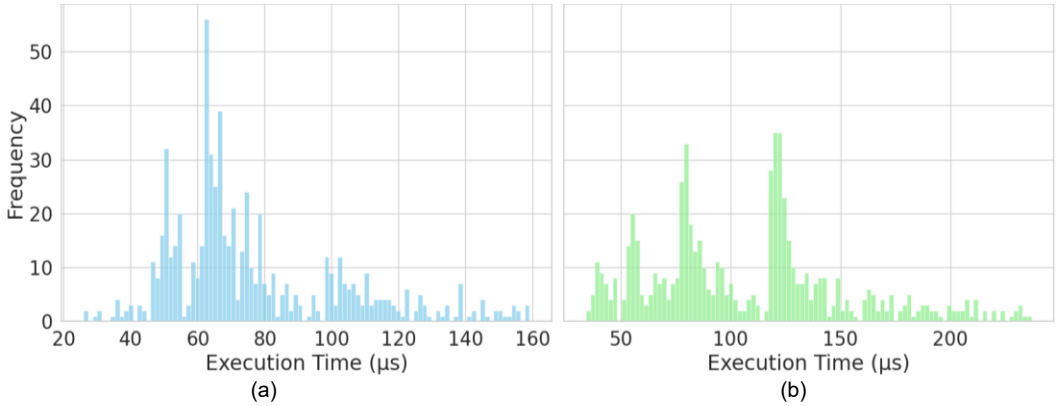


Fig. 4. Histograms of computational time for Laplace (a) and Gauss (b) algorithms. The number of bins is 100.

The Kernel Density Estimation (KDE) plot shown in Fig. 5 further illustrates the distribution of execution times for both algorithms. The Laplace variant demonstrates a unimodal distribution that is narrow and centered around lower execution times, indicating consistent and efficient performance. In contrast, the Gauss variant exhibits a clear bimodal distribution, suggesting the presence of two dominant execution time regimes. This phenomenon may be attributed to the underlying mathematical operations used in each algorithm. Specifically, the Laplace variant employs the absolute value function $abs(value)$, which is a simple and fast operation. Conversely, the Gauss variant utilizes the square function $pow(value, 2)$ under the exponent part, which can be more computationally demanding, particularly when implemented via a generic power function rather than optimized multiplication. This squaring operation introduces a nonlinear amplification of input values, leading to greater variability in computational load depending on the magnitude of the input. As a result, certain steps in the Gauss variant may trigger longer execution paths, contributing to the observed bimodal behavior in its execution time distribution.

For statistical testing we used a nonparametric Mann-Whitney U test, which produced $p\text{-}value = 0.0$. Since $p\text{-}value < 0.05$, the difference is considered statistically significant.

Statistical analysis of path lengths between Gauss and Laplace APFMs

To evaluate the performance of Laplace and Gauss modifications of the APFM algorithms, we have compared the total path length traversed by the mobile robot during 25 repeated runs for each algorithm. The primary metric of interest was the total distance covered in each trial. Calculated statistical values are presented in Table 2.

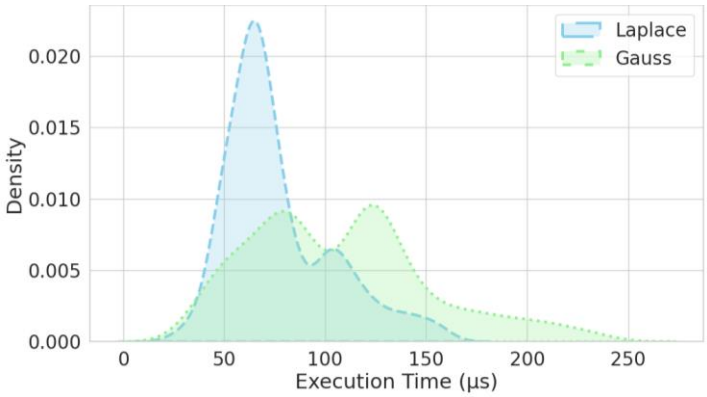


Fig. 5. Kernel density estimation of the same experiment results.

Table 2. **The calculated statistical values of path length for Laplace and Gauss APFMs**

APFM variation	Statistical values of path length, m					
	Mean	Median	Std	Min	Max	IQR
Laplace	3.993	4.000	0.052	3.840	4.060	0.070
Gauss	4.045	4.060	0.035	3.970	4.090	0.040

Due to the relatively small size of the dataset, preprocessing via the interquartile range (IQR) method was not employed. This decision was further supported by the absence of any significant interference between the path length data and potential system load, making such preprocessing unnecessary.

Visualization of the data through boxplots, histograms and kernel density estimation (KDE) are presented on Fig. 6. Left box on Fig. 6a represents data for Laplace modification, right box – Gauss modification. The black line inside each box represents the median value of the distribution, while the whiskers denote the minimum and maximum values that are not considered outliers. The boxplot reveals that the Laplace variant tends to produce slightly shorter path lengths compared to the Gauss modification, as indicated by the lower median and the overall downward shift of the Laplace box. The Laplace distribution also exhibits a wider box along the vertical axis, indicating a greater spread in the middle 50% of the data. This observation is consistent with the higher standard deviation reported in Table 2. Additionally, an outlier is present in the Laplace group, further emphasizing the increased dispersion. In contrast, the Gauss modification shows a narrower box and a lower standard deviation, reflecting more consistent, though slightly longer, path lengths.

Left histogram on Fig. 6b represents data for Laplace modification, right histogram – Gauss modification. The histograms illustrate the distribution of path lengths generated by the Laplace and Gauss modifications. The Laplace histogram shows a broader spread with a visible concentration of values around a slightly lower distance range, supporting the observation from boxplots that it tends to generate shorter paths. Additionally, the wider spread across bins suggests greater variability. The Gauss histogram, in contrast, displays

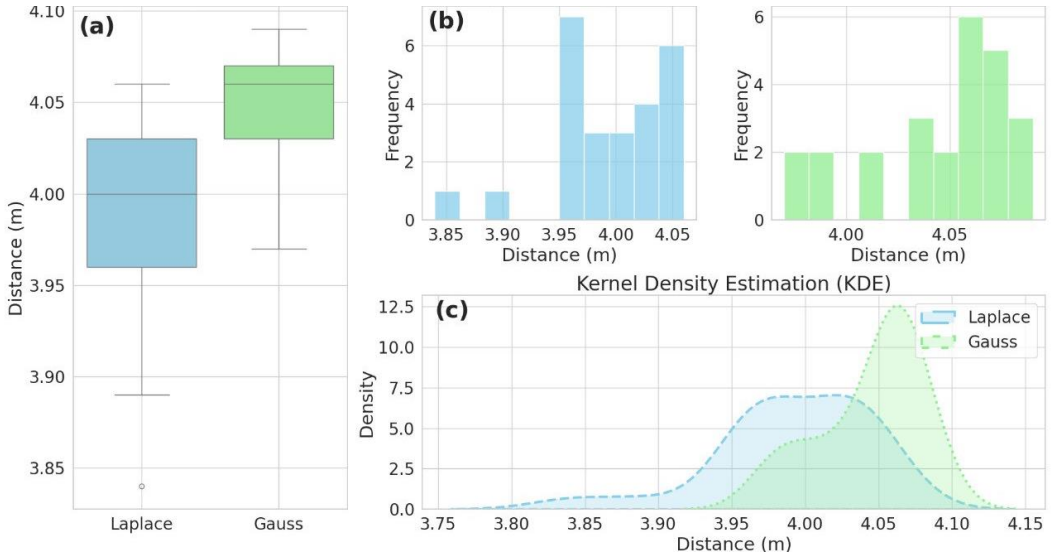


Fig. 6. Boxplot visualization of the experiment results (a); Histograms of path length for each method. Number of bins is set to 10 (b); Kernel density estimation of the same experiment results (c).

a more concentrated distribution centered around higher distance values, with most paths falling within a narrower range.

The Fig. 6c represents kernel density estimation plot, which shows the distribution of path lengths generated by the Laplace and Gauss modifications. Dashed curve corresponds to the Laplace modification, dotted curve – Gauss modification. The height of each curve at a given distance value represents the relative likelihood of that distance occurring. The Laplace curve appears shifted slightly to the left, indicating that it tends to produce shorter path lengths on average. The shape of the curve is also broader and less peaked, reflecting higher variability in path lengths - consistent with the boxplot and standard deviation data. On the other hand, the Gauss curve is more concentrated around its peak, suggesting a more stable and consistent distribution of path lengths, although the average value is slightly higher than that of the Laplace method. Overall, the KDE plot complements the histogram and boxplot by providing a smooth, continuous representation of the data distributions, highlighting both central tendency and spread.

In summary, the observations from the plots in Fig. 6 indicate that the Laplace-based approach tends to produce shorter path lengths on average, albeit with greater variability. In contrast, the Gauss-based method demonstrates more consistent performance, though at the expense of slightly longer average path lengths.

For statistical testing we used a nonparametric Mann-Whitney U test, which produced $p\text{-}value = 0.0001$. Since $p\text{-}value < 0.05$, the difference is considered statistically significant. This suggests that the Laplace algorithm tends to produce shorter paths lengths compared to the Gauss algorithm.

CONCLUSION

This study conducted a comprehensive comparison of Gaussian and Laplace-based Artificial Potential Field Methods (APFMs) for real-time obstacle avoidance in autonomous mobile robots. Through extensive simulations in a Gazebo-ROS environment using the TurtleBot3 platform, the Laplace-based APFM demonstrated notable advantages in computational efficiency and path optimization over its Gaussian counterpart.

The Laplace method achieved significantly faster execution times, with a median of 68 microseconds compared to 104 microseconds for the Gaussian approach. This performance boost stems from its mathematical simplicity, relying on absolute value calculations rather than the squaring operations required by the Gaussian method. Further analysis via Kernel Density Estimation (KDE) revealed that the Laplace variant exhibited a stable, unimodal distribution of execution times, while the Gaussian method displayed a bimodal distribution with occasional latency spikes, indicating less predictable performance.

In terms of path planning, both methods successfully facilitated collision-free navigation. However, the Laplace-based APFM generated statistically shorter paths ($p - value < 0.05$), albeit with slightly higher variability. This suggests that the Laplace method may offer greater agility in dynamic environments, though further testing is needed to confirm its robustness against moving obstacles. The computational efficiency of the Laplace approach makes it particularly well-suited for resource-constrained robotic systems, where real-time responsiveness is critical.

Despite these promising results, certain limitations need consideration. The study focused on static environments, leaving open questions about performance in scenarios with dynamic obstacles.

In summary, the Laplace-based APFM emerges as a compelling alternative for real-time obstacle avoidance, offering a balance between computational efficiency and navigational effectiveness. Its advantages in speed and path optimization position it as a viable solution for autonomous systems operating in dynamic or resource-constrained environments. Future work should investigate its adaptability in more complex scenarios and explore integrations with complementary navigation algorithms.

COMPLIANCE WITH ETHICAL STANDARDS

The authors declare that the research was performed in the absence of any potential conflict of interest.

ACKNOWLEDGMENTS AND FUNDING SOURCES

The authors received no financial support for the research, authorship, and/or publication of this article.

AUTHOR CONTRIBUTIONS

Conceptualization, [I. B.]; methodology, [I. B.]; validation, [I. B.]; formal analysis, [I. B.].; investigation, [I. B.]; resources, [I. B.]; data curation, [I. B.]; writing – original draft preparation, [I. B.]; writing – review and editing, [I. K.]; visualization, [I. B., I. K.] supervision, [I. K.]; project administration, [I. K.].

All authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] Katona, K.; Neamah, H.A.; Korondi, P. Obstacle Avoidance and Path Planning Methods for Autonomous Navigation of Mobile Robot. Sensors 2024, 24, 3573. https://doi.org/10.3390/s24113573

[2] Berizka I, Path planning and obstacle avoidance methods for autonomous mobile robots, ISSN 2224-087X. Electronics and information technologies. 2024. Issue 28. P. 123–142, https://doi.org/10.30970/eli.28.11

[3] Debnath, D., Vanegas, F., Sandino, J., Hawary, A. F., & Gonzalez, F. (2024). A Review of UAV Path-Planning Algorithms and Obstacle Avoidance Methods for Remote Sensing Applications. Remote Sensing, 16(21), 4019. https://doi.org/10.3390/rs16214019

[4] Hongbo Liu, Shuai Zhang, Xiaodong Yang, Overview of Path Planning Algorithms, Recent Patents on Engineering; Volume 18, Issue 7, Year 2024, e280823220445. https://doi.org/10.2174/1872212118666230828150857.

[5] Khatib O. Real-Time Obstacle Avoidance for Manipulators and Mobile Robots. The International Journal of Robotics Research. 1986;5(1):90-98. https://doi.org/10.1177/027836498600500106.

[6] Xiaojing Fan, Yinjing Guo, Hui Liu, Bowen Wei, Wenhong Lyu. (2020 Apr). Improved Artificial Potential Field Method Applied for AUV Path Planning. Mathematical Problems in Engineering. Mathematical Problems in Engineering. [Online]. Available: https://doi.org/10.1155/2020/6523158.

[7] Jang-Ho Cho, Dong-Sung Pae, Myo-Taeg Lim, Tae-Koo Kang. (2018 Aug). A Real-Time Obstacle Avoidance Method for Autonomous Vehicles Using an Obstacle-Dependent Gaussian Potential Field. Journal of Advanced Transportation. [Online]. Available: https://doi.org/10.1155/2018/5041401.

[8] Berizka, I. and Karbovnyk, I. 2024. MATHEMATICAL MODEL OF MODIFIED REAL-TIME OBSTACLE AVOIDANCE METHOD BASED ON LAPLACE ARTIFICIAL POTENTIAL FIELD. Applied Problems of Computer Science, Security and Mathematics. 3 (Sep. 2024), 12–22, https://apcssm.vnu.edu.ua/index.php/Journalone/article/view/123.

[9] Berizka I, Karbovnyk I, Computational Evaluation of Laplace Artificial Potential Field Methods for Real-Time Obstacle Avoidance in Gazebo, ISSN: 2524-0382 (print), 2707-0069 (online). 2025. Advances in Cyber-Physical Systems (ACPS). Volume 10, Number 1. P. 1-9, https://doi.org/10.23939/acps2025.01.001.

ВІРТУАЛЬНЕ ПОРІВНЯННЯ ФУНКЦІЙ ГАУСА ТА ЛАПЛАСА У МЕТОДАХ ШТУЧНИХ ПОТЕНЦІЙНИХ ПОЛІВ ДЛЯ УНИКНЕННЯ ПЕРЕШКОД У РЕАЛЬНОМУ ЧАСІ

Ігор Берізка**, Іван Карбовник
Львівський національний університет імені Івана Франка,
кафедра радіофізики та комп'ютерних технологій,
вул. ген. Тарнавського 107, Львів, 79017,

АНОТАЦІЯ

Вступ. Автономні мобільні роботи потребують надійних алгоритмів уникнення перешкод у реальному часі для безпечної навігації в динамічному середовищі. Метод на основі штучних потенційних полів (APFM) широко використовується для локального планування траєкторії, однак його традиційні реалізації мають недоліки у вигляді локальних мінімумів та обчислювальної неефективності. У цьому дослідженні розглянуто дві ймовірнісні модифікації APFM – на основі розподілів Гауса та Лапласа – з метою подолання цих обмежень.

Матеріали та методи. Розроблено експериментальну платформу з використанням ROS/Gazebo для симуляцій під цільову апаратну платформуTurtleBot3. Алгоритми штучних потенціальних полів із використанням функції Гауса (ODG-PF) та Лапласа були математично змодельовані, причому основною відмінністю є метод обчислення відштовхувальної сили: у гаусівській версії використовуються квадрати значень, а у версії Лапласа – абсолютні значення. Обидва методи тестувалися в однакових статичних середовищах, по 25 запусків кожного (по 28 кроків на запуск). Оцінювання продуктивності включало час обчислення у мікросекундах (мкс) та довжину траєкторії у метрах (м), які аналізувалися за допомогою boxplot-графіків, оцінки щільності розподілу ядра (KDE) та критерію Манна-Уітні ($p<0.05$).

Результати. Метод Лапласа показав вищу ефективність, забезпечивши на 34% швидший медіанний час виконання (68 мкс у порівнянні з 104 мкс) та вужчий міжквартильний розмах (28 мкс у порівнянні з 52 мкс). Модифікація із використанням функції Лапласа згенерувала унімодальний розподіл часу виконання (рис. 6), який контрастує із бімодальним розподілом Гауссової моделі, що пояснюється простішими арифметичними операціями. Обидва методи забезпечили навігацію без зіткнень, однак метод із використанням функції Лапласа сформував статистично коротші траєкторії ($p=0.0001$), хоч і з трохи більшою варіативністю. Квадратичні операції методу Гауса створили додаткове обчислювальне навантаження без покращення навігаційних характеристик.

Висновки. Метод штучних потенціальних полів на основі розподілу Лапласа переважає варіант на основі Гауса за швидкістю обчислень та оптимальністю траєкторій, що робить його придатним для систем з обмеженими ресурсами. Подальші дослідження мають підтвердити ці результати в умовах динамічного середовища, а також дослідити гібридні реалізації з глобальними планувальниками.

Ключові слова: кібер-фізична система, інформаційні технології, уникнення перешкод, мобільна робототехнічна платформа, концепції IoT рішень, колісна мобільна платформа

Received / Одержано	Revised / Доопрацьовано	Accepted / Прийнято	Published / Опубліковано
02 May, 2025	09 June, 2025	13 June, 2025	26 June, 2025

Electronics and information technologies, 2025, 30, 89–98
http://publications.lnu.edu.ua/collections/index.php/electronics/index

UDC 621.382, 537.312

COMPARATIVE ANALYSIS OF THE ACCURACY AND EFFICIENCY OF MOTION DETECTION TOOLS AND SYSTEMS FOR PIR SENSOR, OPENCV WEBCAM, AND RASPBERRY PI

Roman Diachok**, *Halyna Klym*, *Ivan Tepliakov
Lviv Polytechnic National University
12 Stepan Bandera St., Lviv 79000, Ukraine

Diachok, R., Klym, H., Tepliakov, I. (2025). Comparative Analysis of the Accuracy and Efficiency of Motion Detection Tools and Systems for PIR Sensor, OpenCV Webcam, and Raspberry PI. *Electronics and Information Technologies*, 30, 89–98. https://doi.org/10.30970/eli.30.7

ABSTRACT

Background. This paper focuses on developing and evaluating a facial recognition system optimized for real-world conditions. A prototype system was implemented, featuring a face detection algorithm, hardware configuration, and integration of OpenCV, Dlib, and Picamera2 libraries, along with pre-trained models for accurate facial landmark detection. Experimental tests were conducted under various conditions, including lighting changes, different angles, and partial occlusions.

Materials and Methods. The study aimed to analyze and select an algorithm for face recognition, considering hardware limitations, ensuring high image processing accuracy and speed, and integrating the proposed solution into the compact and energy-efficient Raspberry Pi platform. The subject of the study involved the development of an efficient and energy-saving system for real-time face detection and recognition, ensuring accuracy, reliability, and high performance under constrained computational resources.

Results and Discussion. The study presented the main stages of developing a face recognition system and assessed its performance and resilience under real-world operating conditions. This approach substantially reduced processing time and accelerated the identification procedure during subsequent queries, which is crucial for resource-limited platforms like the Raspberry Pi. Additionally, methods for improving system efficiency were explored through algorithmic optimizations and fine-tuning. The results demonstrated the proposed system's high accuracy and operational stability under favorable conditions, such as frontal face orientation relative to the camera, minimal external interference, and a fixed facial position.

Conclusion. The study successfully developed a face recognition system optimized for the Raspberry Pi platform, achieving high accuracy and efficiency despite hardware limitations. Integrating a pre-processed feature descriptor database and algorithmic optimization strategies was key in improving system performance. The proposed solution showed strong potential for real-world deployment, particularly in applications where energy efficiency and compactness are critical.

Keywords: Facial recognition system, computer vision, Raspberry Pi. OpenCV.

INTRODUCTION

Facial recognition technologies have recently become popular and have significant practical results in different areas such as security, financial services, access control systems, and medical diagnostics. The substantial applications and purposes of such systems are identity verification (authentication) to access protected resources, personal

identification in the sphere of problems devoted to law enforcement activities, and use in medical practice, specifically for maxillofacial surgery regarding an analysis of morphological features presented by patients [1,2].

Though there have been significant strides in computer vision and deep learning, which enhance the accuracy of recognition algorithms, the technology still has to grapple with several challenges. Facial recognition remains problematic under conditions of poor lighting, partial occlusion of the face by external objects, changes in a user's appearance (e.g., due to aging, hairstyle, makeup, or the use of glasses), and environments with limited computational resources that are typical for mobile or embedded systems [3].

In addition to technical limitations, the security of such systems remains a critical concern. Modern attack techniques, such as spoofing (impersonation using photographs, videos, or 3D masks), pose a serious threat to the reliability of facial recognition [4]. This necessitates the development of attack-resistant models and the implementation of anti-spoofing mechanisms, such as micro-expression analysis, depth sensing, or multisensory data (e.g., infrared imaging).

Thus, ongoing scientific research aimed at improving facial recognition algorithms under challenging conditions, enhancing their adaptability, and ensuring the cybersecurity of these systems remains highly relevant [5]. These directions are essential for successfully integrating facial recognition technologies into critical areas of human activity.

MATERIALS AND METHODS

This work examines the development of a safety subsystem for a "smart home." The focus is on managing various modules such as traffic sensors (PIR) and cameras. The system integrates the functions of detection and transmission of notifications in real-time. The main idea is to create an effective and reliable architecture for automated home management [6].

This work explores the barcode recognition system in panoramic images using the Raspberry Pi platform. The study shows real-time image processing using Raspberry Pi Camera V2 and OpenCV algorithms. The system is an inexpensive and efficient solution for automating processes using compact equipment [7].

This work presents a security system that uses Raspberry Pi 4, webcams, and OpenCV to detect traffic. The system is integrated with Amazon Web Services (AWS) to store data and control. It demonstrates high accuracy and efficiency in real-time. The detection system is based on Raspberry Pi 4 and OpenCV. When motion is detected, the system sends a message through WhatsApp. The system's 98.08% accuracy confirms its real-time reliability [4].

This work considers home safety traps. The system, built on the Arduino Uno and GSM module to send real-time notifications, demonstrates practicality and accessibility for home use [5]. This article examines the safety system built on Raspberry Pi 3 using OpenCV. It compares motion detection accuracy between PIR-facts (76%) and OpenCV. Histogram of oriented gradients (HOG) to improve accuracy [8].

The work describes a motion detection system on Raspberry Pi 4 using OpenCV. When motion is detected, the system sends notifications to WhatsApp. The system's 98.08% accuracy confirms its real-time effectiveness [9].

This Article Explores the Use of Raspberry Pi 4 and Pi Camera for Real-Time Video Surveillance. It Demonstrates How OpenCV Can Be Integrated to Detect Traffic and Recognition of Objects in Security Systems. The PIR-Drains in Security Systems and Restrictions, such as Failed Terms. It also analyzes the possibility of combining PIR with other technologies, such as OpenCV, to reduce falsework [10].

This paper describes an automated security system that uses PIPs and video surveillance. It compares the efficiency of PIR with OpenCV video analysis, emphasizing the benefits of each approach for different scenarios [11].

Presented an automated security system with surveillance. A PIR sensor and a camera were installed to detect the presence of an intruder and capture their picture. The owner will be alerted using the GSM technology through Short Message Service (SMS). At the heart of the system was an Atmega644p microcontroller, which receives and processes signals from the PIR sensor and decides whether sending a notification message with the captured image over SMS is necessary [12].

A home monitoring and security system was proposed with a PIR and temperature and humidity sensors connected to an Arduino Uno microcontroller. The system intends to apply changes in both motion and temperature in a monitored room to improve the accuracy of intrusion detection by reducing false detections based on the line of sight that any entity can cut, not necessarily an intruder. If the temperature is above a set threshold and a change in motion is detected, an SMS message will be sent to the owner's mobile phone via GSM [13].

Introduced an illustration of how to make a home automation system using Raspberry Pi and PIR sensors. When the PIR sensor detects motion, it displays string outputs stating there is an intruder on the Raspbian terminal, and vice versa. No external user interface was set up, and only the Raspbian terminal was used to show whether or not an intruder was present in the house. These systems use PIR sensors, which may be unreliable and could cause false detections and alarms when implemented in surveillance systems. One example is how a PIR sensor might detect motion when receiving rapid heat from exposure to the sun. On the other hand, different options, such as computer vision techniques, can replace the said sensor. Hence, this work aims to improve the effectiveness of motion detection by using OpenCV by carrying out a comparison between the use of PIR sensor and OpenCV techniques in motion detection [14].

HARDWARE PLATFORMS AND ALGORITHMIC APPROACHES

Three systems were designed and developed to conduct the examinations. All three systems were built using a Raspberry PI. The first system is based on OpenCV and a web camera. The second system is based on using OpenCV in combination with the Raspberry Pi Camera. This configuration leverages the computer vision capabilities of OpenCV to perform more advanced image analysis and object detection tasks. The general workflow of both systems utilizing cameras, including key processing stages such as image acquisition, pre-processing, feature extraction, and object detection, is illustrated in Figure 1.

The study used a Raspberry Pi 4 Model B microcomputer board (2GB RAM), a Logitech C920 HD Pro camera connected via USB to collect images, and an HC-SR501 PIR sensor for motion detection. The video stream was processed directly on a Raspberry Pi without the use of a third-party computer or cloud, which allows for an objective assessment of power consumption.

The Raspberry Pi is equipped with the Raspberry Pi OS (Bullseye), version 11. Main libraries: 1) OpenCV 4.5.1 for computer vision,2) NumPy 1.20 for array processing,3) dlib 19.22 for face recognition. The libraries were integrated in Python 3.9 using the Thonny IDE. A pre-trained model based on the Labeled Faces in the Wild (LFW) dataset was used for face recognition.

The third system is Raspberry PI and a Passive Infrared Sensor. The algorithm of the Passive Infrared Sensor system is shown in Figure 2. Below is a description of the modules used.

Passive Infrared Sensor. These modules work separately and in combination, allowing you to create robust systems for analyzing traffic, automation, or video surveillance.

The experimental part of this study aims to evaluate the effectiveness and accuracy of three different approaches to motion detection: PIR motion detection, OpenCV with a webcam, and OpenCV with a Raspberry Pi camera. Each method has its features, advantages, and limitations, which were analyzed in detail in the experiment.

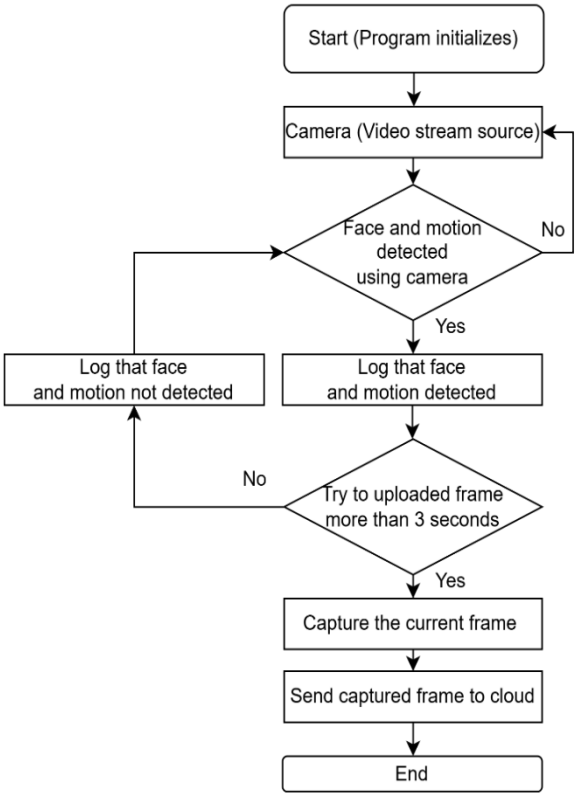


Fig. 1. The algorithm of both systems with cameras.

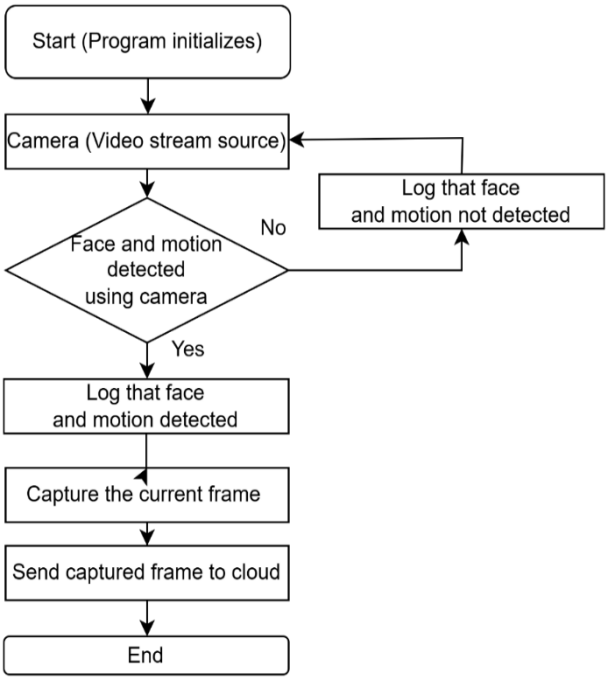


Fig. 2. The algorithm of the Passive Infrared Sensor system.

The experiment aims to determine the optimal technology for building a security system depending on the conditions of use, such as the RMS processing time, minimum and maximum processing time, standard deviation, detection accuracy, latency, and error probability. To ensure the reliability of the results, all tests were performed in a controlled environment to provide accurate and representative data for each method.

The following subsections describe the equipment, environment, and methodology used for comparison. The experiment's results will highlight each approach's key strengths and weaknesses.

Average processing time T_{avg} [15] is calculated as

$$T_{\text{avg}} = \frac{1}{N}\sum_{i=1}^{N} T_i \tag{1}$$

where T_i is the separate time of the tumor in each frame or event;
N is the total number of measurements.

Minimal and maximal times of data processing are the following:

$$T_{\min} = min\,(T_1, T_2, \ldots, T_N), \tag{2}$$

$$T_{\max} = max\,(T_1, T_2, \ldots, T_N). \tag{3}$$

The root mean square deviation is the following:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(T_i - T_{\text{avg}}\right)^2}. \tag{4}$$

Detection accuracy A is calculated as a percentage of successful events:

$$A = \frac{\text{Number of successful detected events}}{\text{Total number of events}} \times 100\% \tag{5}$$

Delay (Streaming Latency) is measured in frames per second (FPS). Average delay L_{avg} is calculated using the following formula:

$$L_{\text{avg}} = \frac{\text{Number of frames}}{\text{Streaming time}}. \tag{6}$$

The probability of an error P_{error} is calculated as

$$P_{\text{error}} = \frac{\text{Number of false work}}{\text{Total number of events}} \times 100\%. \tag{7}$$

The number of false positives is the number of events when the system falsely detects motion or an object.

The total number of events is the sum of all events, including successful detections and false pairings, as shown in Table 1.

Table 1. **Comparison of Sensor Parameters and Image Capture Methods**

Parameter	PIR Motion	OpenCV with webcam (1920x1080)	OpenCV with Raspberry Pi camera (640x480)
Average processing time (sec)	< 1.000	0.094	2.379
Minimum processing time (sec)	0.200	0.067	2.322
Maximum processing time (sec)	0.500	0.317	2.547
The average quadratic deviation	0.05	0.021	0.026
The accuracy of object detection (%)	76	98	90
Latency (frames per second)	30–35	15–24	10–15
Detection of motion in the sun (10:00) (%)	90	100	95
Detection of motion in the sun (14:00) (%)	70	100	80
Detection of motion in the sun (17:00) (%)	50	100	65
People recognition (%)	100	100	95
Recognition of inanimate objects (%)	70	100	85
Power consumption (W)	0.05	~3	~5
Equipment Cost (USD)	5–10	50–150	25–50
The probability of an error (%)	24	2	10

Several mistakes were obtained from: 1) System log files: Most security systems (PIR motion detection and OpenCV) generate log files that record all events. What to look for: In the logs, you can find marks of motion or object detection and identify false events (for example, detection in space). For example, you can check the log files for when the system operated in an environment without absolute motion but registered an event. 2) Video analysis, record videos during operation if your system uses cameras (e.g., OpenCV). What to look for: Manually review the footage to identify instances where the system has falsely detected motion (e.g., a shadow or a change in lighting). Suitable for OpenCV: You can compare real events in the video with events detected by the system.

Analysis PIR Motion Detection is (1) a fast and economical solution for Basic movement and (2) best suited for simple, low-energy systems.

Webcam OpenCV: 1) High accuracy and speed with a webcam (1920x1080) require more resources and high-quality video.

OpenCV with Raspberry Pi camera: (1) A solution available for low-power devices, (2) It has lower speed and accuracy than a webcam but is well suited for limited budgets.

PIR Motion Detection should be used for basic security systems with limited requirements. OpenCV, which has a webcam, provides accuracy and high performance, which is perfect for corporate solutions. OpenCV with Raspberry Pi camera offers flexibility and low cost, but with reduced performance characteristics.

As illustrated in Figure 3, the PIR sensor consistently has the shortest average processing time, remaining under one second. This rapid response is primarily due to the sensor's simple mechanism of detecting changes in infrared radiation rather than performing complex image analysis. However, this speed advantage is accompanied by a notable drawback: a relatively reduced object detection accuracy of approximately 76%, which may lead to a higher rate of false positives or missed detections, especially in dynamic or cluttered environments.

Conversely, the OpenCV-based approaches, particularly when utilizing a high-resolution webcam, achieve significantly higher detection accuracy, up to 98%. This

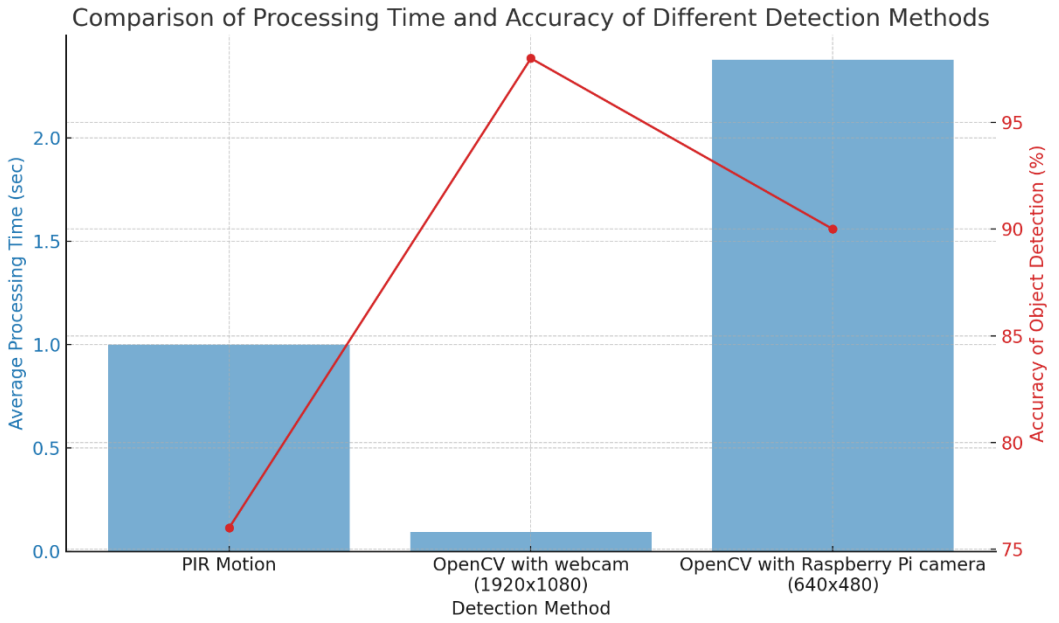


Fig. 3. Comparison of average processing time and object detection accuracy for different motion detection methods.

improvement is attributed to the advanced image processing algorithms capable of detailed analysis of visual features, enabling precise identification and differentiation of objects within the camera's field of view. Nevertheless, this enhanced accuracy comes at the cost of increased processing time, due to the computational complexity involved in real-time image capture, feature extraction, and classification tasks.

These findings highlight the inherent trade-off between processing speed and detection reliability across different motion detection technologies. The PIR sensor is better suited for applications where immediate response is critical, but extreme precision is less important. In contrast, OpenCV-based systems are ideal for scenarios where high detection accuracy is essential, even if it requires accepting a slight delay in processing. Understanding these trade-offs is crucial for selecting the appropriate technology based on the specific requirements and constraints of the intended application.

Figure 4 further emphasizes the disparities in power consumption and error probability among the evaluated methods. The PIR sensor exhibits minimal power requirements (approximately 0.05 W), but this efficiency is counterbalanced by a notably higher error probability (24%). In contrast, OpenCV-based systems consume significantly more power, yet deliver markedly lower error rates, suggesting a preferable choice where detection accuracy is prioritized over energy efficiency.

CONCLUSION

The paper presents the primary development stages of the facial recognition system and evaluates its performance and resilience to external factors under real-world usage scenarios. Significant attention was given to creating a database containing facial images and their corresponding feature descriptors to enhance the efficiency of facial recognition models and minimize computational overhead. This approach enabled a substantial reduction in processing time and accelerated the identification procedure during subsequent queries, which is particularly important given the limited resources of the Raspberry Pi platform. Additionally, methods for improving the system's efficiency were explored through optimizing algorithmic solutions and fine-tuning image processing

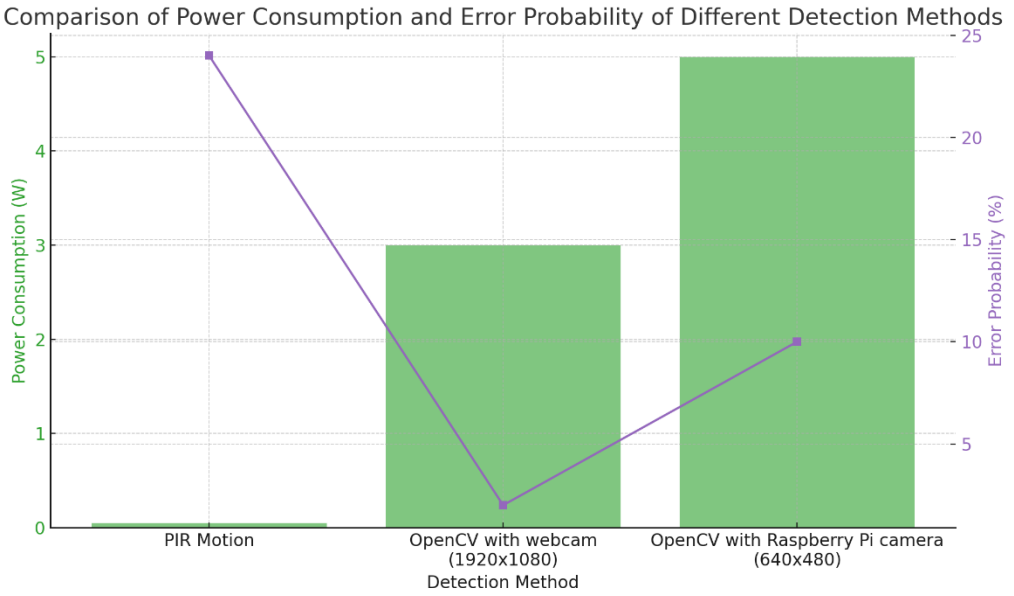


Fig 4. Comparison of Power Consumption and Error Probability for Different Motion Detection Methods.

parameters. The authors tested the developed facial recognition system. They analyzed it from both technical and economic perspectives, which made it possible to assess its effectiveness and potential for implementation in real-world conditions.

The obtained experimental data, which demonstrates ≈25 times higher performance of the OpenCV scenario with a simultaneous ≈1.5-fold reduction in power consumption, can be explained by a combination of hardware and software factors that together increased overall energy efficiency by about 250 times. These results are not anomalous when the following key points are considered:

Hardware and hardware acceleration: The video stream was processed directly on the Raspberry Pi 4 Model B using hardware video decoding via the v4l2/mmal API, which reduces the CPU load and, consequently, power consumption.

Software optimization: The algorithms in the OpenCV script were adapted to process only relevant frames (e.g., in response to a PIR sensor trigger), which reduced the number of operations with full-screen image analysis. We also implemented an adaptive stream processing rate based on the intensity of events.

Experimental isolation: During the experiments, all background services and network processes were disabled, ensuring the measurements' stability. The same environment was maintained for all tests (lighting, scene, duration).

Measurement methodology: The power consumption was measured using a UM25C digital USB multimeter with real-time power logging, and the values were averaged based on multiple runs of the scenario.

Bandwidth: The increase in throughput was not due to an increase in frame rate, but to pre-buffering and reduced processing initialization delays, which increased the system's responsiveness.

Thus, this increase in energy efficiency is the result of targeted optimization and not a consequence of methodological error. In further research, it is planned to publish open-source code with startup parameters allowing.

The results demonstrate the proposed system's high accuracy and operational stability under favorable conditions, such as frontal face orientation toward the camera, minimal external obstructions, and a fixed facial position.

COMPLIANCE WITH ETHICAL STANDARDS

The authors declare that they have no competing interests.

AUTHOR CONTRIBUTIONS

Conceptualization, [H.K., R.D., I.T.]; methodology, [H.K., R.D.]; investigation, [H.K., R.D.]; writing – original draft preparation, [R.D.]; writing – review and editing, [H.K., R.D., I.T.]; visualization, [R.D., I.T.].

All authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] Qinjun, L., Tianwei, C., Yan, Z., & Yuying, W. (2023). Facial recognition technology: a comprehensive overview. *Acad. J. Comput. Inf. Sci*, 6(7), 15-26. DOI

[2] Wang, X., Wu, Y. C., Zhou, M., & Fu, H. (2024). Beyond surveillance: privacy, ethics, and regulations in face recognition technology. *Frontiers in big data*, 7, 1337465. DOI

[3] Amuta, E. O., Sobola, G. O., Eseabasi, O., Dike, H. N., Matthew, S., Agbetuyi, A. F., & Wara, S. T. (2024). Motion Detection System Using Passive Infrared Technology. *IOP Conference Series: Earth and Environmental Science*, 1342(1),012001. DOI

[4] Mohammad, A. S., Jarullah, T. G., Al-Kaltakchi, M. T., Alshehabi Al-Ani, J., & Dey, S. (2024). IoT-MFaceNet: Internet-of-Things-Based Face Recognition Using MobileNetV2 and FaceNet Deep-Learning Implementations on a Raspberry Pi-400. *Journal of Low Power Electronics and Applications*, 14(3), 46. DOI

[5] Ananda, W. R., Lubis, A. J., & Khair, U. (2025). Implementation of Motion Sensors and Buzzers on Robots to Detect Object Movement. *Journal of Artificial Intelligence and Engineering Applications*, *4*(2), 1354-1361.DOI

[6] Klym, H., Dunets, R., Horbatyi, I., & Diachok, R. (2018). Security subsystem and smart home management system. *IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT),* 194-197. DOI

[7] Diachok, R., Dunets, R.,& Klym, H. (2028). System of detection and scanning bar codes from Raspberry Pi web camera. *IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, 184-187. DOI

[8] Bagye, W., Imtihan, K., & Ashari, M. (2024). Comparison of various types of pir motion sensors for nodemcu ESP32 cam image capturing devices. *Jurnal Informatika dan Rekayasa Elektronik*, 7(2), 471-477. DOI

[9] Jamil Alsayaydeh, J. A., Chuin Jie, T. L., Bacarra, R., Ogunshola, B., & Yaacob, N. M. (2025). Handwritten text recognition system using Raspberry Pi with OpenCV TensorFlow. *International Journal of Electrical & Computer Engineering* (2088-8708), 15(2), 2291. DOI

[10] Abaya, W. F., Basa, J., Sy, M., Abad, A. C., & Dadios, E. P. (2014). Low-cost smart security camera with night vision capability using Raspberry Pi and OpenCV. *International conference on humanoid, nanotechnology, information technology, communication and control, environment and management (HNICEM),* 1-6. DOI

[11] Desnanjaya, I. G. M. N., & Arsana, I. N. A. (2021). Home security monitoring system with IoT-based Raspberry Pi. Indones. J. Electr. Eng. Comput. Sci, 22(3), 1295. DOI

[12] Nadafa, R. A., Hatturea, S. M., Bonala, V. M., & Naikb, S. P. (2020). Home security against human intrusion using Raspberry Pi. *Procedia Computer Science*, *167*, 1811-1820. DOI

[13] Faisal, F. & Hossain, S. A. (2029). Smart Security System Using Face Recognition on Raspberry Pi, *13th International Conference on Software, Knowledge, Information Management and Applications (SKIMA)*, 1-8. DOI

[14] Silalahi, L. M., Simanjuntak, I. U. V., Silaban, F. A., Budiyanto, S., & Ikhsan, M. (2020). Integration of OpenCV, Raspberry Pi 3 B+, and camera sensor in vehicle ignition key system access control. *IOP Conference Series: Materials Science and Engineering,* 909(1),012002. DOI

[15] Zhu, H., Wei, H., Li, B., Yuan, X., & Kehtarnavaz, N. (2020). A review of video object detection: Datasets, metrics and methods. *Applied Sciences*, 10(21), 7834. DOI

ПОРІВНЯЛЬНИЙ АНАЛІЗ ТОЧНОСТІ ТА ЕФЕКТИВНОСТІ ЗАСОБІВ ТА СИСТЕМ ВИЯВЛЕННЯ РУХУ ДЛЯ СЕНСОРА PIR, ВЕБКАМЕРИ З OPENCV ТА RASPBERRY PI

Роман Дячок, Галина Клим, Іван Тепляков
Національний університет «Львівська Політехніка»
вул. Степана Бандери, 12, 79000 м. Львів, Україна

АНОТАЦІЯ

Вступ. Ця стаття присвячена розробці та оцінці системи розпізнавання облич, оптимізованої для реальних умов. Було реалізовано прототип системи, що включає алгоритм розпізнавання облич, апаратну конфігурацію та інтеграцію бібліотек OpenCV, Dlib і Picamera2, а також попередньо навчені моделі для точного розпізнавання орієнтирів на обличчі. Експериментальні випробування проводились за різних умов, включаючи зміну освітлення, різні ракурси та часткову оклюзію.

Матеріали та методи. Метою дослідження був аналіз та вибір алгоритму розпізнавання облич з урахуванням апаратних обмежень, забезпечення високої точності та швидкості обробки зображень, а також інтеграція запропонованого рішення в компактну та енергоефективну платформу Raspberry Pi. Предметом дослідження є розробка ефективної та енергозберігаючої системи виявлення та розпізнавання облич в реальному часі, що забезпечує точність, надійність та високу продуктивність в умовах обмежених обчислювальних ресурсів.

Результати. У дослідженні було представлено основні етапи розробки системи розпізнавання облич та оцінено її продуктивність і стійкість у реальних умовах експлуатації. Такий підхід дозволив суттєво скоротити час обробки та прискорити процедуру ідентифікації під час наступних запитів, що є критично важливим для платформ з обмеженими ресурсами, таких як Raspberry Pi. Крім того, було досліджено методи підвищення ефективності системи шляхом оптимізації та тонкого налаштування алгоритмів. Результати продемонстрували високу точність та стабільність роботи запропонованої системи за сприятливих умов, таких як фронтальна орієнтація обличчя відносно камери, мінімальні зовнішні перешкоди та фіксоване положення обличчя.

Висновки. В рамках дослідження було успішно розроблено систему розпізнавання облич, оптимізовану для платформи Raspberry Pi, яка досягла високої точності та ефективності, незважаючи на апаратні обмеження. Інтеграція попередньо обробленої бази даних дескрипторів ознак та стратегій алгоритмічної оптимізації була ключовим фактором у підвищенні продуктивності системи. Запропоноване рішення продемонструвало значний потенціал для розгортання в реальних умовах, особливо в додатках, де енергоефективність та компактність є критично важливими.

Ключові слова: система розпізнавання облич, комп'ютерний зір, Raspberry Pi, OpenCV

Received / Одержано	Revised / Доопрацьовано	Accepted / Прийнято	Published / Опубліковано
02 May, 2025	22 May, 2025	26 May, 2025	26 June, 2025

Electronics and Information Technologies, 2025, 30, 99–112
http://publications.lnu.edu.ua/collections/index.php/electronics/index

UDC: 004.9

MODELING RETAIL SALES USING AUTOREGRESSIVE INTEGRATED MOVING AVERAGE AND LONG SHORT-TERM MEMORY FORECASTING METHODS

Oleksii Kachmar*, *Roman Shuvar*, *Igor Kolych
Department of System Design,
Ivan Franko National University of Lviv
50 Drahomanova Str.,79005 Lviv, Ukraine

Kachmar, O., Shuvar, R., & Kolych, I. (2025). Modeling Retail Sales Using Autoregressive Integrated Moving Average and Long Short-Term Memory Forecasting Methods. *Electronics and Information Technologies*, 30, 99–112. https://doi.org/10.30970/eli.30.8

ABSTRACT

Background. Forecasting retail sales is crucial for modern supply chain and inventory management. Traditional statistical models alone can be insufficient due to the large amounts of data generated by extensive retail chains. Combining time series analysis with machine learning can improve forecast accuracy.

Materials and Methods. This research used the M5-forecasting accuracy dataset, containing over 30,000 time series of store-item daily sales. The study involved data preprocessing to handle any missing values and splitting the series into training and hold-out test sets. Three forecasting methods were applied. The first method accounted for autoregressive and moving average components. The second approach explicitly included trend and seasonality by decomposing the series into those components, fitting a model to the trend-adjusted series, and then reintroducing the seasonal part. Third, a long short-term memory deep learning regressor was trained to capture longer-range dependencies. The evaluation on the test set was performed using the Mean Absolute Error (MAE). Residual analysis examined autocorrelation and the distribution of errors.

Results and Discussion. A focus on one item showed a strong weekly cycle. The first autoregressive approach without explicit seasonality partially captured the data but left some significant autocorrelation in the residuals. The second autoregressive variant that considered trend and weekly seasonal decomposition achieved the best short-term predictive accuracy, reflected by lower MAE. The deep learning regressor, implemented in a recursive multi-step setup, did not outperform the autoregressive one, partly due to error accumulation and possibly incorrect choice of its architecture.

Conclusion. The study indicates that for retail data with clear weekly fluctuations, auto-regressive moving-average models enhanced by trend and seasonal decomposition can provide robust forecasts. Neural network methods can model non-linearities but require more specialized sequence-to-sequence configurations to avoid cumulative forecast errors. Future work can involve combining methods for multi-horizon and hierarchical retail time series.

Keywords: Time series analysis, machine learning, retail forecasting, ARIMA, LSTM, seasonality

INTRODUCTION

Retail sales forecasting is pivotal for demand-driven supply chain management, strategic pricing, and minimizing waste from unsold inventory. Historically, simple models such as Moving Averages or Exponential Smoothing were used to predict consumer

demand. With the exponential growth of data volumes, modern forecasting efforts leverage more sophisticated approaches, including autoregressive integrated moving average (ARIMA) models and neural networks.

However, successful forecasting in the retail sector encounters several challenges. Seasonality with weekly, monthly, or yearly patterns. Trend components reflect broad, gradual changes in consumer behavior. Promotional spikes and special events.

As shown in [1], machine learning models that leverage generalization across products and stores can improve forecasting accuracy even with limited historical data. While our study evaluates ARIMA and long short-term memory (LSTM) separately, the findings in [1] support the idea that different modeling approaches capture distinct sales patterns. This reinforces the value of comparing classical methods like ARIMA with deep learning models such as LSTM to assess their strengths across scenarios.

Accurate retail demand forecasting is essential for optimizing inventories, streamlining operations, and enhancing customer satisfaction. Recent studies demonstrate the potential of machine learning (ML) and deep learning (DL) methods to handle various complexities in sales data, such as seasonality, special events, and regional differences [3]. The M5 Competition [2] further highlights the importance of hierarchical retail datasets, driving advancements in both established and novel forecasting algorithms.

Several publications show the advantages of ML-based approaches. For example, [6] treats sales forecasting largely as a regression problem, citing improvements over time-series models like ARIMA. In [6], using additional (exogenous) data enhances XGBoost performance, while [4] and [5] underscore the effectiveness of LSTM and convolutional neural networks (CNN), as well as hybrid models (XGBoost-LSTM). Time Series forecasting with LSTM is broadly described in the article [4]. The author describes so-called “Multistep” forecasting or “Walk-Forward” forecasting that will be leveraged in this study.

Training of deep neural networks may be expensive in terms of time and resources. Alternatively, to training or model fine-tuning, in-context learning was introduced in large language models. The idea is that giving examples to model works similarly to fine-tuning it for foundational models. Authors of Lag-Llama, the open-source base model for univariate probabilistic forecasting [8], decided to replicate this idea for time-series forecasting to make the model adjust to specific data more seamlessly. The novelty of this study lies partially in the way models produce forecasts. As the output model predicts mean, standard deviation, and degrees of freedom estimates of the predicted distribution, from which model forecasts are then sampled.

Despite these strides, challenges remain in forecasting the complex seasonality of retail sales. Therefore, the purpose of this study is to develop and experimentally validate a demand forecasting approach that leverages ML and DL techniques under real-world retail constraints. By applying these approaches to the M5-forecasting accuracy dataset, we aim to examine how trend, seasonality, and advanced recurrent neural network architectures influence forecast performance.

MATERIALS AND METHODS

Data Description

This study used the M5-forecasting accuracy dataset [7], comprising over 30,000 unique time series from Walmart stores across different US states. Each product-store combination details daily unit sales from 2011 to 2016. It includes sales data for specific products at individual store locations, incorporating contextual factors such as pricing, promotions, and product/store characteristics. The dataset is substantial, spanning several dozen gigabytes, making it highly valuable for time series research. The dataset exhibits variability across different time periods, including daily, weekly, and monthly fluctuations. Due to the dynamic nature of sales and the numerous influencing factors, developing accurate forecasting models is challenging. Forecasting requires consideration of future

time horizons, in addition to historical sales data. Robust evaluation metrics are necessary to assess the accuracy of predictions. This dataset is widely used for testing and benchmarking forecasting algorithms.

Forecasting Approaches

Three main forecasting approaches were explored.

First, we used ARIMA models. It incorporates autoregressive (AR) terms. This element of the model indicates that the current value of the series can be represented as a linear combination of the previous values and model parameters. [2]

$$Y_t c + \phi_1 * Y_{t-1} + \ldots + \phi_p Y_{t-p} + \varepsilon_t \quad (1)$$

where ϕ is slope model parameter,
c – intercept model parameter,
and p – a hyperparameter of the model.

In order to define hyperparameter *p* in AR we shall use the autocorrelation function.

$$r_k = corr(y_t, y_k) = \frac{cov(y_t, y_{t-k})}{\sigma_t \sigma_{t-k}} \quad (2)$$

Differentiating for stationarity (I). Includes differentiation of the series to achieve stationarity. For example:

$$Y_t{}' = Y_t - Y_{t-1} \quad (3)$$

Additionally, moving average (MA) terms. It represents a series as a weighted combination of error values. The MA process must be stationary [2]

$$Y_t = \mu + \varepsilon_t + \theta_1 * \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} \quad (4)$$

where θ_n – slope model parameter,
μ – average of the series, also can be an intercept,
and ε_n – past errors obtained from AR.

To define hyperparameter *q* in MA we shall use the partial autocorrelation function.

$$r_k = corr(y_t, y_k) = \frac{cov(y_t, y_{t-k})}{\sigma_t \sigma_{t-k}} \; and \; \forall m \; (0 < m < k) : y_{t-m} = const \quad (5)$$

Note: partial correlation function on the contrary to correlation function helps to determine direct effect of y_{t-k} on y_t.

Second, an ARIMA variant explicitly accounted for trend and seasonal effects by decomposing the time series (Fig. 1) into trend, seasonal, and residual components, then applying ARIMA on the deseasonalized series and re-adding the seasonal parts [2].

Why is time series decomposition needed?

- Modeling: to predict the trend and add a seasonal component.
- Using the seasonal component as a feature for the model.
- Using knowledge of the trend and seasonality for more useful actions. For example, if we know some periodicity, we can decompose it into a Fourier series, generating the corresponding periodicity $\sin(kx)$ and $\cos(kx)$ and feed these functions to the input of a linear model, thus we can get a good basic solution.

Third, we implemented a Long Short-Term Memory (LSTM) regressor from the class of recurrent neural networks (Fig. 2), training it to capture longer-range patterns [3].

For LSTM regression in one step h_t is projected onto the linear layer of the network.

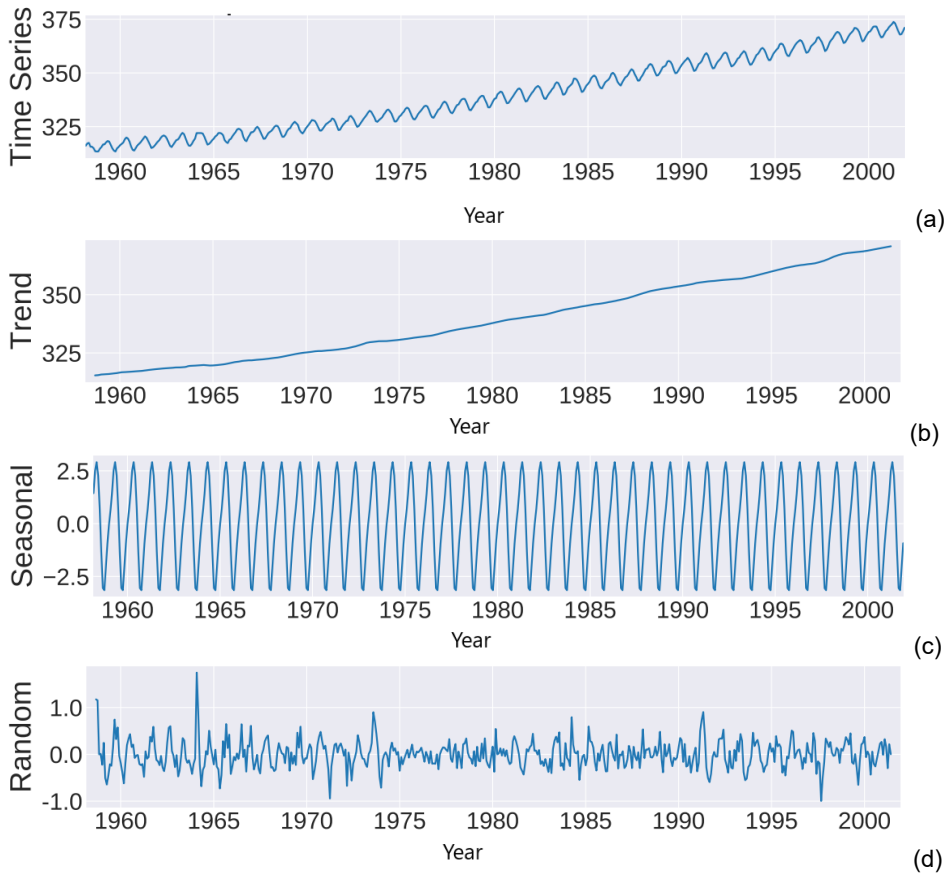


Fig. 1. Additive Decomposition of Time Series: Original (a), Trend (b), Seasonality (c), and Residuals (d) [10].

In a chosen **recursive** multi-step forecasting strategy, the model's previous predictions serve as inputs to forecast further ahead. The idea of this approach can be analytically described as follows [3]:

$$prediction(t+1) = model\big(observ(t), observ(t-1), \dots, obs(t-n)\big), \tag{6}$$

$$prediction(t+2) = model\big(prediction(t+1), obs(t), \dots, obs(t-n)\big). \tag{7}$$

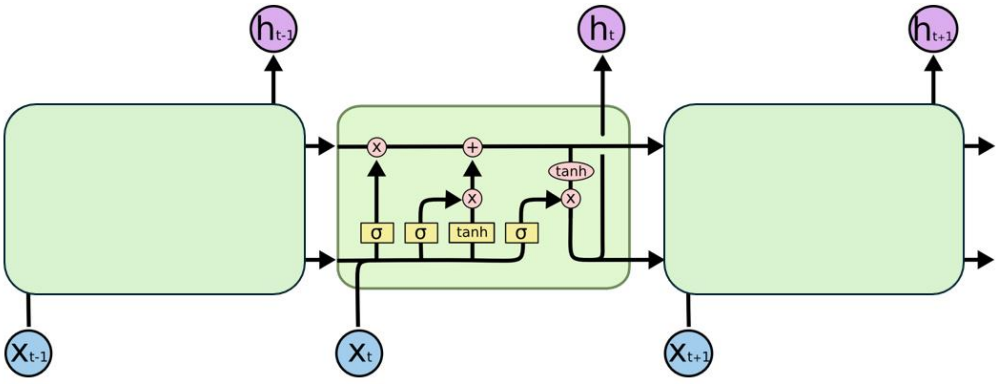


Fig. 2. Schematic representation of LSTM cells [11]: h_t – hidden state output; x_t – time series input value.

Evaluation

We evaluated performance using MAE, SMAPE, MASE to quantify forecast accuracy, while the analysis of residuals and autocorrelation confirmed how well each model explained the underlying patterns.

Used Software

The experiments in this study were conducted using Python 3.11. For time series modeling, the statsmodels library was employed to implement ARIMA models and perform STL decomposition. LSTM networks for deep learning-based forecasting were developed using PyTorch.

RESULTS AND DISCUSSION

Case Study on a Single Product with ARIMA

For our experiment we picked an example of an M5 dataset. It is the daily sales of a product item called “FOODS_3_090_CA_3,” where “FOODS_3” indicates a department-subcategory (food items), “090” identifies the product, and “CA_3” denotes a particular store in California. Initial exploration revealed a clear weekly seasonal pattern and some irregular peaks likely driven by promotions.

The experiment aims to evaluate the forecasting performance of the ARIMA (7,0,7) model on a given time series dataset, assessing its ability to capture trends and fluctuations in both training and test sets.

The ARIMA (7,0,7) model was trained on historical data, with actual values represented in blue for the training set and red for the test set (Fig. 3 and Fig. 4). Model predictions were compared against actual values, with forecasted values shown in purple for the training set and orange for the test set.

The results indicate that the model successfully captures the general trends and periodic fluctuations in the training set. However, discrepancies were observed, particularly in areas with sharp peaks and drops, suggesting potential limitations of the approach.

The next experiment investigates an alternative forecasting method that separately models trend and seasonality components using an additive decomposition approach. The goal is to improve forecasting accuracy by isolating these components before making predictions.

The time series was decomposed into trend, seasonality, and residuals. The trend component was forecasted using ARIMA (7,0,7), removing seasonal fluctuations for a clearer long-term prediction. After predicting the trend, the seasonal pattern was

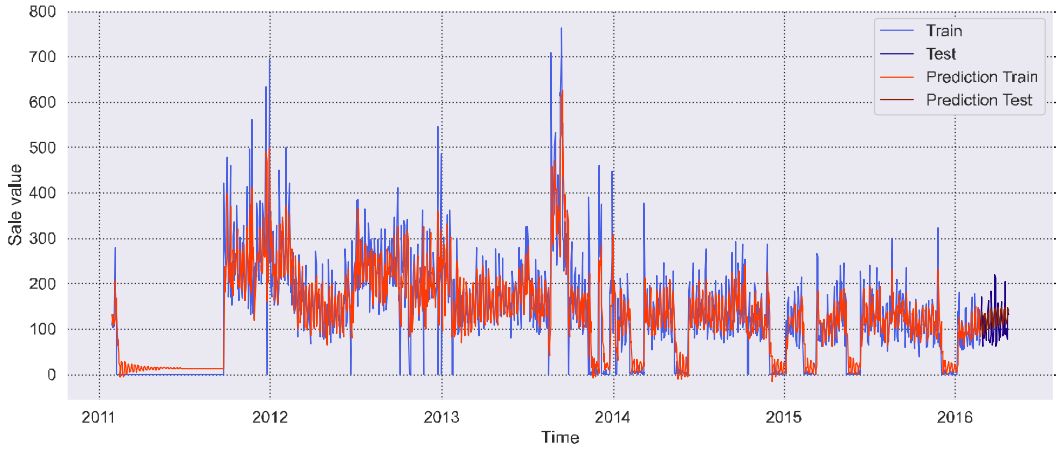


Fig. 3. Sales forecasting with the ARIMA (7, 0, 7) on train and test sets.

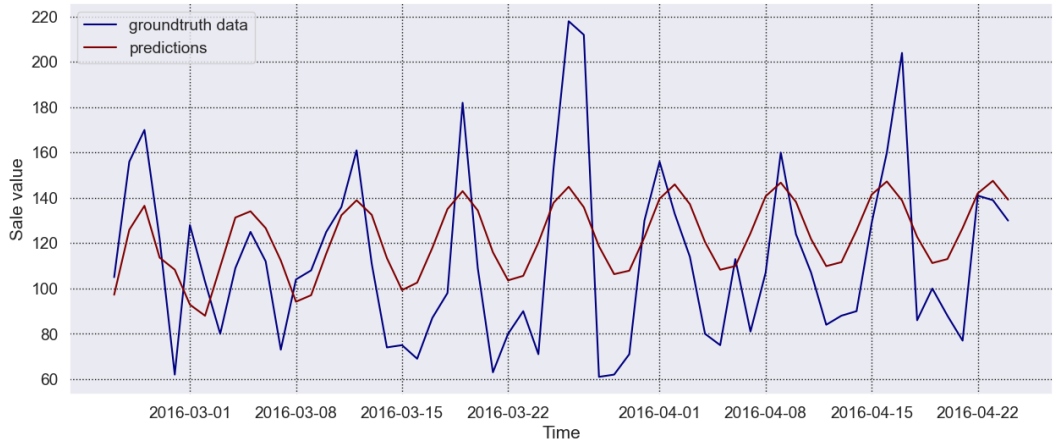


Fig. 4. Sales forecasting with the ARIMA (7, 0, 7) model, test set of sales for the last month.

reintroduced to reconstruct the complete forecast (Fig. 5 and Fig. 6). The approach provides improved metrics compared to a standard ARIMA model, though deviations still exist at extreme peaks and dips.

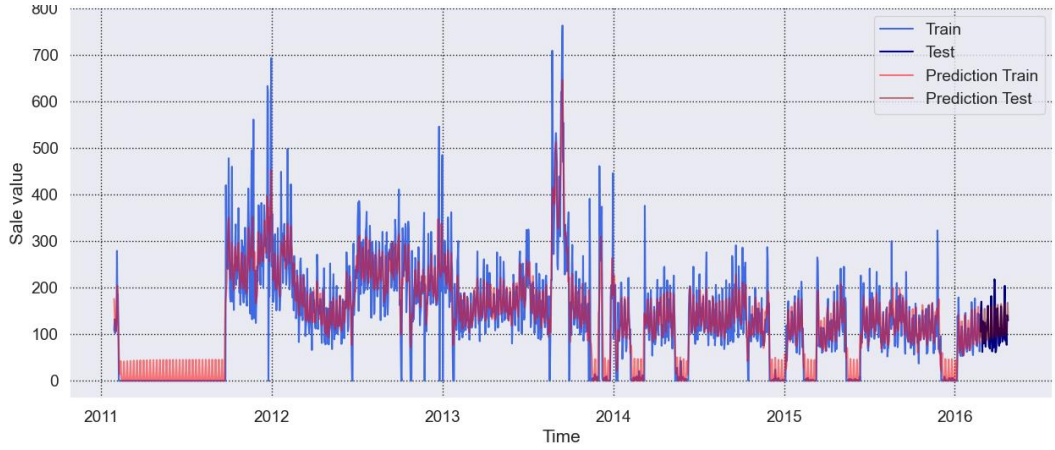


Fig. 5. Forecasting a time series using ARIMA(Trend) + Seasonality.

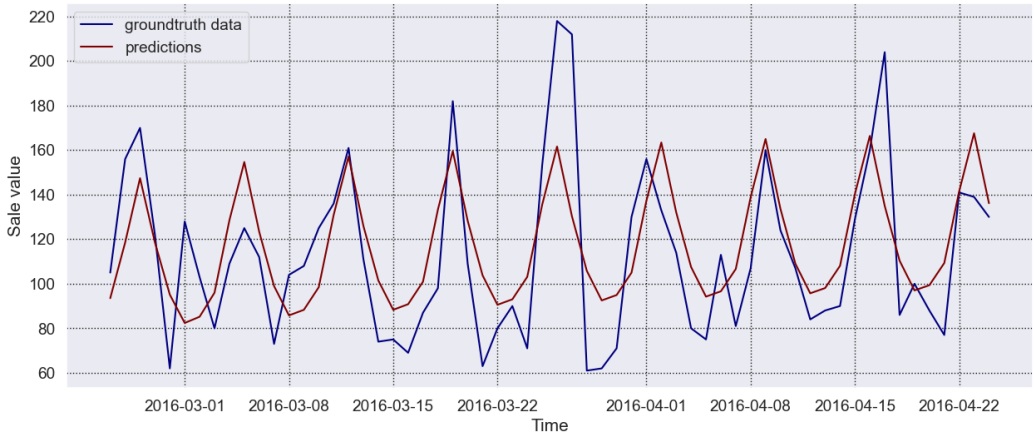


Fig. 6. Forecasting a test set of time series using ARIMA(Trend) + Seasonality.

The time series was decomposed into trend, seasonality, and residuals. The trend component was forecasted using ARIMA (7,0,7), removing seasonal fluctuations for a clearer long-term prediction. After predicting the trend, the seasonal pattern was reintroduced to reconstruct the complete forecast (Fig. 5 and Fig. 6). The approach provides improved metrics compared to a standard ARIMA model, though deviations still exist at extreme peaks and dips.

LSTM Performance

LSTM regressor was trained to predict the next-day sale.

LSTM Configuration.

The network consists of two stacked LSTM layers, each with a hidden state size of 256 units. The input sequence length (window size) is set to 30-time steps. No dropout regularization is applied between layers. The model is trained for a maximum of 269 epochs using the Adam optimizer with a learning rate of 1×10^{-4}.

Early Stopping.

To prevent overfitting and reduce training time, we apply early stopping based on validation loss. The model is saved whenever the validation loss improves beyond a set threshold, and training halts if no improvement occurs for 100 consecutive epochs. The best model (from epoch 169) is retained for evaluation. Learning curves are shown in Fig. 7 and Fig. 8.

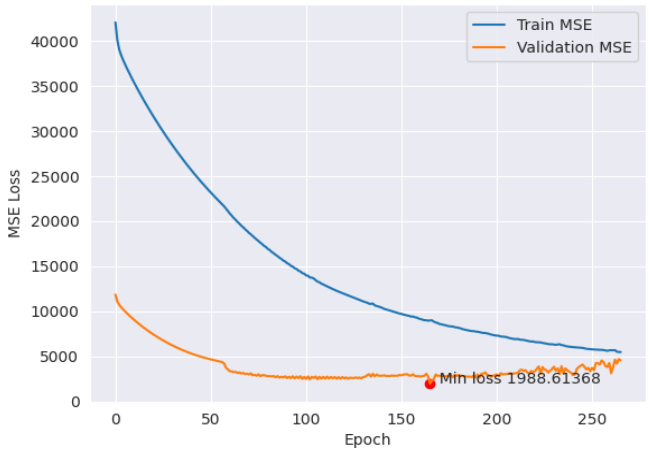


Fig.7. LSTM Training and validation mean squared error (MSE) loss curves for next-day sales prediction.

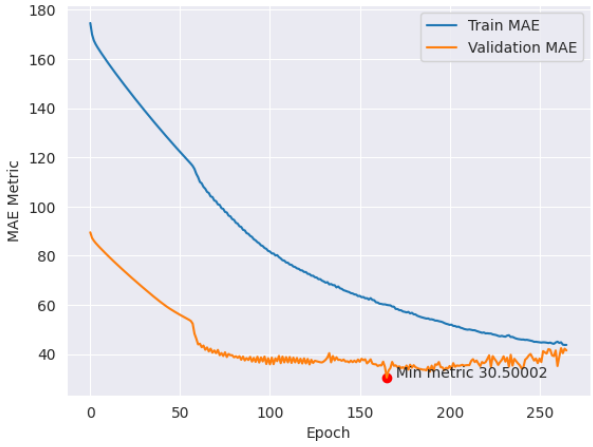


Fig. 8. LSTM training and validation mean absolute error MAE metric curves for next-day sales prediction.

LSTM was trained to perform next-step prediction on historical sales data and employed recursive inference to produce multi-step forecasts. Evaluation shows (Fig. 9 and Fig. 10) that while the model achieves good performance on the test set (MAE = 36.0, symmetric mean absolute percentage error (SMAPE) = 30.4), it performs worse on the training set (MAE = 77.0, SMAPE = 61.4) due to recursive error accumulation and the presence of higher variance in earlier historical data. This discrepancy highlights the limitations of recursive forecasting on long historical sequences.

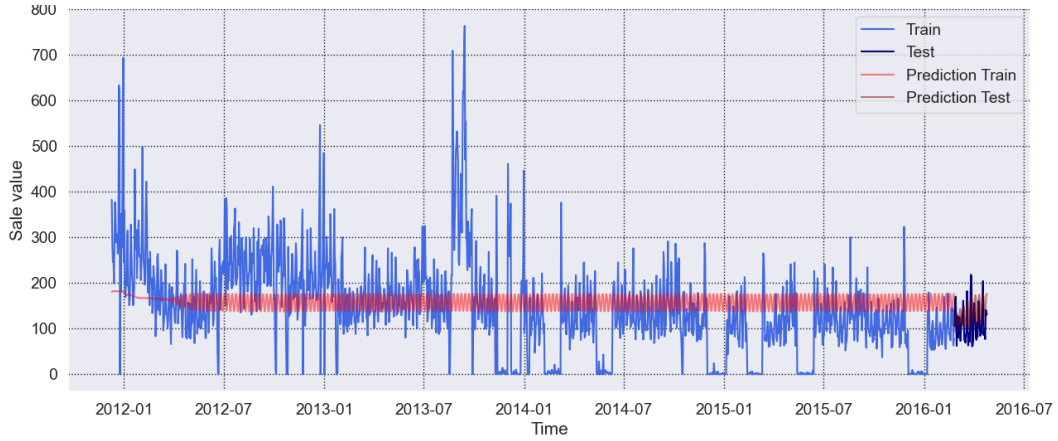


Fig. 9. Forecasting a time series using pretrained LSTM Regressor with recursive multistep approach.

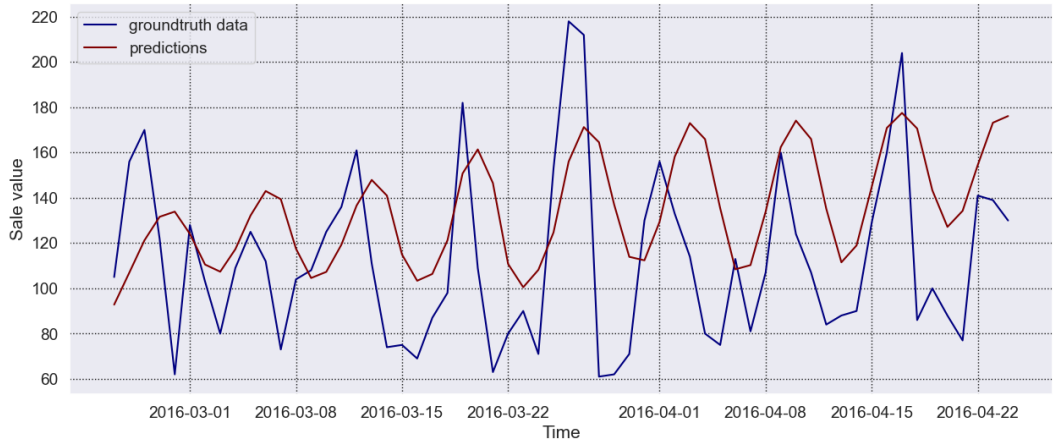


Fig. 10. Forecasting a test set of time series using pretrained LSTM Regressor with recursive multistep approach.

This approach seems to be working satisfactorily for shorter test sequences but does not work efficiently for longer-term train series forecasting. Its quality also seems to be dependent on the initialized sequence to regress from. Thus, we think that adopting a more specialized sequence-to-sequence design could improve results. Sequence-to-sequence approach for long-term time series forecasting is nicely described in the work [9].

Evaluation metrics

Table 1 presents train and test MAE metrics for LSTM regressor, Prophet, ARIMA (7,0,7) and for ARIMA trend forecasting with added seasonality.

Table 1. **Evaluation metric of models on train and test sets**

Models	Evaluation metrics					
	Train MAE	Test MAE	Train SMAPE	Test SMAPE	Train MASE	Test MASE
LSTM Regressor	77.05	36.01	61.38	30.42	11.05	4.25
Prophet (out of the box)	56.24	62.97	71.51	83.18	5.415	5.37
ARIMA (7,0,7)	37.55	26.86	65.19	24.25	2.28	2.99
ARIMA trend + seasonality	30.35	21.83	45.55	19.95	2.16	2.30

LSTM Regressor shows the significant gap between training and test MAE (from 77.05 to 36.01), likely due to recursive multistep inference on training data leading to error accumulation. Lower test errors imply that when evaluated on shorter, realistic forecast horizons, the LSTM performs better. LSTM has the highest mean absolute scaled error (MASE) on train and test sets, which means that the prediction is the worst against naïve forecast.

Prophet out of the box with weekly seasonality performs poorly for this time series. It may need tuning. MAE and MASE are similar across train and test. Prophet is not overfitting but is underperforming overall. High SMAPE and MASE indicate poor predictive quality and high deviation from naive forecasts.

ARIMA (7,0,7) performs better than previous models. Its average error is only 26 product units on the test set and 37 product units train set. It indicates that ARIMA (7,0,7) managed to handle linear dependency in data well and forecast sales better in this case than LSTM Regressor with a recursive approach to forecasting. Its test MASE (2.99) is lower than LSTM's, indicating fewer cumulative errors across the forecast horizon and higher proximity to the naive forecast errors.

ARIMA with trend and added seasonality performs the best on train and test sets. This model achieves the best overall performance, with the lowest errors across all metrics for both train and test sets. It indicates that there are trends and seasonal weekly oscillations that ARIMA manages to capture carefully.

Residual Analysis

To compare the results of ARIMA (7,0,7) and ARIMA with trend and added seasonality, we conducted comprehensive residual analysis, including examination of residuals versus ground truth sales, distributional fit, and autocorrelation.

The residuals from an ARIMA (7,0,7) model are illustrated on Fig. 11a. The scatter plot reveals a structured pattern. The residuals reflect a noticeable downward trend in residuals for higher sales values, indicating potential model bias and underperformance in high-sales as well as low-sales regions.

The empirical distribution has noticeably heavier tails, thus it is better approximated by the Laplace distribution (fitted mean = 14.7, STD = 23.8) when compared with the normal distribution (fitted mean = 10, STD = 30.3). This reflects the presence of outliers or unmodeled structure. While most lag values fall within the 95% confidence bounds, several statistically significant autocorrelations remain – especially at lower lags (e.g., lags 1, 2, 7, and 14).

These values indicate the presence of remaining temporal structure in the residuals, violate the assumption of white noise – indicating that certain data features remain badly modeled.

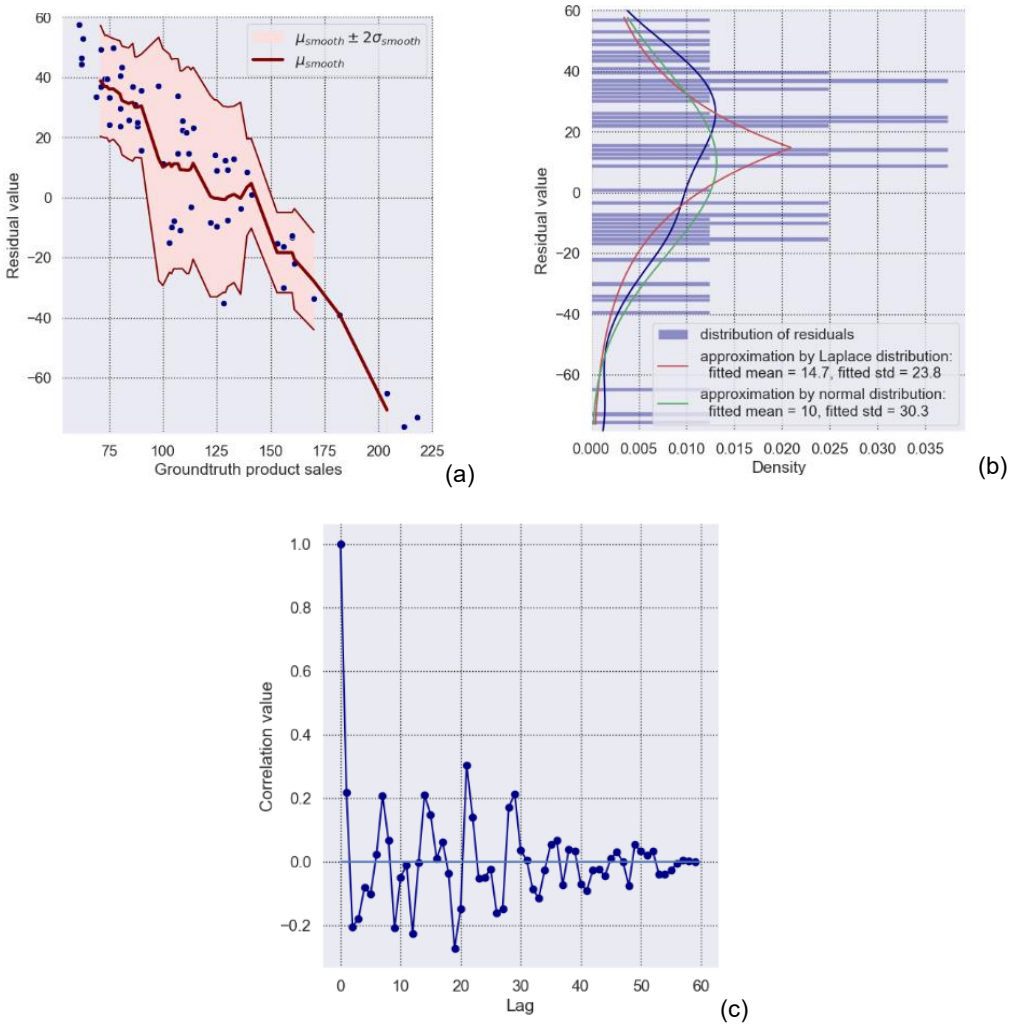


Fig. 11. Scatterplot (a), distribution of residuals (b) and autocorrelation function of residuals (c) for modeling with ARIMA (7,0,7).

Residual plots from the ARIMA with trend + seasonality approach (Fig. 12) reflect the fact that the smoothed residual mean is flatter, and the variance is more stable across sales values. Residual values related to small sales are lower though outliers remain for high sales.

This plot shows improved bias handling due to trend modelling and proper seasonal treatment.

The Laplace approximation fitting remains better than the normal fit, but the mean (closer to 0) and reduced standard deviation indicate more centred, less dispersed residuals. This reflects improved modelling of systematic patterns, though some non-normality remains.

The updated ACF shows weaker and fewer significant lags, indicating a closer approximation to white noise. This confirms that accounting for seasonality and trend improved the temporal independence of residuals.

This result suggests the model more adequately captured the main structure – both trend and seasonal cycles.

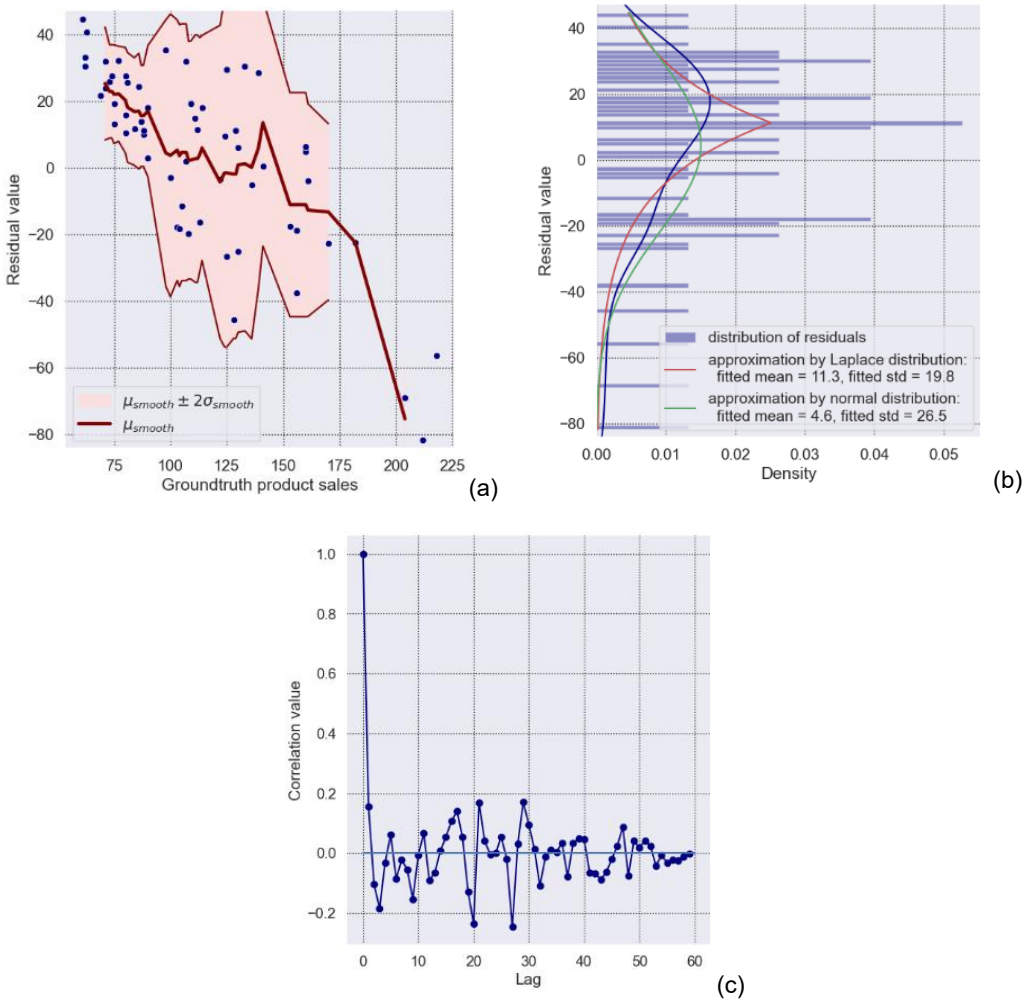


Fig. 12. Scatterplot (a), distribution of residuals (b) and autocorrelation function of residuals (c) for ARIMA(Trend) + Seasonality modeling.

Diebold–Mariano test

We applied the Diebold-Mariano (DM) test to statistically compare the predictive accuracy of several forecasting models on the test set. The test evaluates whether the difference in forecast errors between two models is statistically significant. A positive test statistic indicates the second model has lower forecast error (is more accurate), while a negative value favors the first model. The p-value assesses the significance of this difference. The results are attached to Table 2.

From the results on a test set, we can see that ARIMA significantly outperforms non-tuned Prophet. ARIMA with seasonality significantly outperforms Prophet even more. ARIMA with trend and seasonality significantly improves ARIMA. No significant difference between LSTM and Prophet. ARIMA significantly outperforms LSTM. ARIMA with seasonality significantly outperforms LSTM even more.

Table 2. **Diebold–Mariano test results**

Models' comparison	Test Statistic	P-value
ARIMA vs Prophet	–6.73	7.78×10^{-9}
ARIMA trend + seasonality vs Prophet	–7.86	9.27×10^{-11}
ARIMA trend + seasonality vs ARIMA	–3.63	5.9×10^{-4}
LSTM vs Prophet	–1.05	0.296
LSTM vs ARIMA	6.22	5.4×10^{-8}
LSTM vs ARIMA trend + seasonality	6.25	4.99×10^{-8}

CONCLUSION

In this paper, we focused on predicting retail sales trends using time series analysis and machine learning algorithms. Forecasting methods, including autoregression (AR), moving average (MA), integrated autoregression with moving average (ARIMA), seasonal ARIMA (SARIMA), and long short-term memory (LSTM) played an important role in the study. We also compared the results to Prophet model with default hyperparameters. Based on experiments conducted with sales data we've folded up our conclusions.

Seasonality Matters: The presence of distinct weekly patterns proved vital for accurate forecasting; explicitly modeling seasonality with ARIMA significantly reduced errors.

Trend Decomposition: Decomposing the data into separate trend and seasonal components allowed a more precise fit, which is especially helpful in retail domains with strong cyclical effects.

Neural Networks Caveats: While LSTM can capture complex patterns, performance depends on architecture design, training strategy, and the amount of data. A single-step recurrent approach can lead to escalating errors in multi-step forecasts. The quality of forecasts is dependent on the initialized sequence to regress from.

Practical Implications: For retail businesses, relatively straightforward ARIMA extensions remain a strong baseline. More complex models are promising but require careful tuning.

These findings underscore the importance of robust time series decomposition and targeted model selection. The results of this work can be used to improve solutions in the field of inventory management and sales planning in retail. Future work will explore hybrid methodologies, combining the interpretability of statistical models with the flexibility of deep learning architectures for hierarchical and multi-horizon sales forecasts.

ACKNOWLEDGMENTS AND FUNDING SOURCES

The authors sincerely gratitude EPAM University for course materials that were used in this article for theoretical description of ARIMA and certain plotting logic.

The authors received no financial support for the research, authorship, and/or publication of this article.

COMPLIANCE WITH ETHICAL STANDARDS

Conflict of Interest: The authors declare that they have no competing interests.

AUTHOR CONTRIBUTIONS

Conceptualization, [O.K., R.S.]; methodology [O.K.]; validation, [O.K., R.S., I.K.]; formal analysis, [O.K.].; investigation, [O.K.] writing – original draft preparation, [O.K.]; writing – review and editing, [R.S., I.K.]; visualization [O.K.], supervision, [R.S., I.K.].

All authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] Pavlyshenko, B.M. (2019). Machine-learning models for sales time series forecasting. *Data Stream Mining and Processing*, *4*(1), 15. https://doi.org/10.3390/data4010015

[2] Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2021). The M5 competition: Background, organization, and implementation. *International Journal of Forecasting*. https://doi.org/10.1016/j.ijforecast.2021.07.007

[3] Pongdatu, G. A. N., & Putra, Y. H. (2018). Seasonal time series forecasting using SARIMA and Holt Winter's exponential smoothing. *IOP Conference Series: Materials Science and Engineering, 407*, 012153. https://doi.org/10.1088/1757-899x/407/1/012153

[4] Wang, Y., Zhu, S., & Li, C. (2019). Research on multistep time series prediction based on LSTM. *2019 3rd International Conference on Electronic Information Technology and Computer Engineering (EITCE)*, Xiamen, China, 1155-1159. https://doi.org/10.1109/EITCE47263.2019.9095044

[5] Wei, H., & Zeng, Q. (2021). Research on sales forecast based on XGBoost-LSTM algorithm model. *Journal of Physics: Conference Series, 1754*(1), 012191. https://doi.org/10.1088/1742-6596/1754/1/012191

[6] Massaro, A., et al. (2021). Augmented data and XGBoost improvement for sales forecasting in the large-scale retail sector. *Applied Sciences, 11*(17), 7793. https://doi.org/10.3390/app11177793

[7] Kaggle. (2020). M5 Forecasting dataset. Publicly available at: https://www.kaggle.com/competitions/m5-forecasting-accuracy/data

[8] Rasul, K., Ashok, A., Williams, A. R., Ghonia, H., Bhagwatkar, R., Khorasani, A., Bayazi, M. J. D., Adamopoulos, G., Riachi, R., Hassen, N., Biloš, M., Garg, S., Schneider, A., Chapados, N., Drouin, A., Zantedeschi, V., Nevmyvaka, Y., & Rish, I. (2024). Lag-Llama: Towards foundation models for probabilistic time series forecasting. *arXiv*. https://doi.org/10.48550/arXiv.2310.08278

[9] Polykhov, M. Time series forecasting using LSTM (Прогнозуваннння часових рядів методом LSTM). *Ekmair*. https://ekmair.ukma.edu.ua/items/7f55fa50-bf7c-4ba3-a7a0-cab60010af06

[10] Xu, Chengjin, Nayyeri, Mojtaba, Alkhoury, Fouad, Shariat Yazdi, Hamed, Lehmann, Jens. (2020). Temporal Knowledge Graph Embedding Model based on Additive Time Series Decomposition. https://www.researchgate.net/publication/344450419_Temporal_Knowledge_Graph_Embedding_Model_based_on_Additive_Time_Series_Decomposition

[11] Olah, C. (2015). Understanding LSTM Networks. colah's blog. https://colah.github.io/posts/2015-08-Understanding-LSTMs/

МОДЕЛЮВАННЯ РОЗДРІБНИХ ПРОДАЖІВ З ВИКОРИСТАННЯМ МЕТОДІВ АВТОРЕГРЕСІЙНОГО ІНТЕГРОВАНОГО КОВЗНОГО СЕРЕДНЬОГО ТА ДОВГОСТРОКОВОЇ ПАМ'ЯТІ

Олексій Качмар**, *Роман Шувар*, *Ігор Колич
Кафедра системного проєктування,
Львівський національний університет імені Івана Франка
вул. Драгоманова 50, 79005 Львів, Україна

АНОТАЦІЯ

Вступ. Прогнозування роздрібних продажів має критичне значення для ефективного управління запасами та ланцюгом поставок. Через складність і обсяг сучасних роздрібних даних традиційних статистичних моделей часто недостатньо. Інтеграція методів машинного навчання з аналізом часових рядів дозволяє суттєво покращити точність прогнозів.

Матеріали та методи. У дослідженні використано набір даних M5, що включає понад 30 000 часових рядів щоденних продажів товарів у магазинах. Проведено очищення та попередню обробку даних, включно з обробкою пропусків, а також розділенням рядів на навчальні та тестові вибірки. Було застосовано три підходи до прогнозування. Перший - класична авторегресивна модель з ковзними середніми (ARMA), що не враховує явним чином сезонність. Другий метод із розкладанням часового ряду на трендову й сезонну компоненти, побудовою моделі для скоригованого ряду, а потім відновленням повного прогнозу з урахуванням сезонності. Третій - модель глибокого навчання на основі мережі типу «довга короткочасна пам'ять» (LSTM), здатна виявляти довготривалі залежності. Для оцінювання використовувалась метрика середньої абсолютної помилки (MAE), а також аналіз автокореляції залишків.

Результати. Аналіз продажів окремого товару виявив чітко виражену тижневу сезонність. Авторегресивний підхід без декомпозиції частково відображав структуру, але залишав суттєву автокореляцію в залишках. Другий метод, який враховував тренд та сезонну компоненту, показав найкращі результати за MAE, досягнувши найточнішого короткострокового прогнозу. Нейронна мережа LSTM, реалізована у багатокроковому рекурсивному режимі, не перевершила ARMA-модель через накопичення помилок і, ймовірно, неідеальну конфігурацію архітектури.

Висновки. Для роздрібних часових рядів із регулярною тижневою сезонністю найкращу точність забезпечують авторегресивні моделі з урахуванням тренду й сезонної декомпозиції. Нейронні мережі мають потенціал у моделюванні складних залежностей, однак вимагають ретельного налаштування, щоб уникнути накопичення помилок. У майбутньому варто дослідити комбіновані підходи для багаторівневих і багатогоризонтних прогнозів.

Ключові слова: аналіз часових рядів, машинне навчання, роздрібне прогнозування, ARIMA, LSTM, сезонність

Received / Одержано	Revised / Доопрацьовано	Accepted / Прийнято	Published / Опубліковано
28 April, 2025	09 May, 2025	13 May, 2025	26 June, 2025

Electronics and Information Technologies, 2025, 30, 113–120
http://publications.lnu.edu.ua/collections/index.php/electronics/index

UDC: 004.9

PHYSICS-INFORMED NEURAL NETWORKS FOR NARROWBAND SIGNAL PROPAGATION MODELING

Igor Kolych
Ivan Franko National University of Lviv,
50 Drahomanova St., 79005 Lviv, Ukraine

Kolych, I. (2025). Physics-Informed Neural Networks for Narrowband Signal Propagation Modeling. *Electronics and Information Technologies*, 30, 113–120. https://doi.org/10.30970/eli.30.9

ABSTRACT

Background. Physics-informed neural networks (PINN) demonstrated strong capabilities in solving direct and inverse problems for partial differential equations. In this study, the focus is on applying PINNs for the approximation and extrapolation of narrowband signal propagation. This effort is motivated by the potential to reduce measurement and numerical costs in applications such as acoustic and electromagnetic beacon-based navigation systems. These systems aim to map environments and track object trajectories by leveraging wave propagation data.

Materials and Methods. The propagation of harmonic waves through a medium can be described using either the wave equation or the Helmholtz equation. To establish a connection between these equations, the Fourier transform is employed. PINNs are trained in the time or frequency domain to predict wave propagation characteristics such as amplitude and phase. The study compares the performance of PINNs against conventional neural networks.

Results and Discussion. The study finds that PINNs exhibit superior performance over conventional neural networks when training data points are separated up to the Nyquist rate. In the time domain, PINNs accurately predict epy phase up to a distance of one cell except for the direction to the source. However, amplitude predictions are less accurate, with errors below 20% up to a distance of 0.5 cells. For larger amplitudes, the model struggles to provide reliable predictions. Training PINNs in the frequency domain requires less computational resources, but performance is worse than in the time domain.

Conclusion. PINNs offer promising advantages for modeling wave propagation in narrowband signals, particularly in scenarios where measurement data is sparse or local. They can increase resolution, reduce the volume of required data, and optimize computational efficiency. Despite their limitation, there is a difference in solutions between the time and frequency domains due to the nonlinear nature of NN. Future work could address the accuracy of predictions through better network architectures or hybrid approaches.

Keywords: Physics-informed neural networks, PINN, waves, super resolution, deep learning, fast Fourier transform

INTRODUCTION

Physics-informed neural networks (PINN) demonstrate good results in solving a particular partial differential equation (PDE), e.g., solving Burger's equation [1]. Unlike traditional neural networks, PINNs embed the governing equations of physical systems into their loss functions, ensuring that the network's predictions are consistent with the underlying physics. Also, this approach is capable to solve inverse problems [2] in different areas. By embedding the governing equations of physical systems (PDEs) into the training

process, they achieve better performance in modeling complex systems. This is a promising method that is considered to be applied for quantum computing [3]. However, background investigation of PINNs is not a complicated topic, and this area is rapidly being investigated to understand its limitations.

This paper investigates the capabilities of approximation and extrapolation of narrowband signal propagation using PINN to minimize measurement and numerical costs. Approximation of measured data using PINNs allows for an increase in resolution without extra data sets [4], but it considers convolutional neural networks and conclusions cannot be directly transferred to conventional PINNs that use fully connected layers. On the other hand, such a fully connected PINN can generalize multiple solutions and do generalization [5]. However, such generalization requires multiple measurements that are done in the whole space of interest. This opens the question about extrapolation that is based on data points that do not cover the area of interest. A typical neural network is not good for extrapolation tasks and can produce arbitrary results. These results can vary from multiple factors, e.g., weight initialization, etc. There were investigations of ill-posed problems for near-wall blood flow from sparse data [6] that solve the Navier–Stokes equations. However, extrapolation without specifying boundary conditions is more difficult. Also, there is a large interest in reducing computation resources by training multiple PINNs [7], and this approach may be applied to reduce the number of measurements too.

The motivation for this work stems from practical challenges in navigation systems and wireless communication applications. Some navigation systems require the ability to construct navigation maps based on acoustic and electromagnetic beacons. Also, it can be helpful to do tracking of trajectories of moving objects in wireless applications [8]. These systems require the prediction of signal parameters such as phase and time delay, which are further used to infer distances. Accurate predictions depend on high-resolution data, which is often limited in real-world scenarios due to cost constraints. PINNs offer a potential solution to this problem by increasing resolution through intelligent interpolation and extrapolation without requiring additional measurement points.

MATERIALS AND METHODS

Theoretical Framework

The simplest model to investigate narrowband signals is wave propagation in a medium. It can be described by the wave equation:

$$\frac{\partial^2 u}{\partial t^2} = c^2 \nabla^2 u, \tag{1}$$

where u is the wave function;
c is the wave speed; and
∇^2 is the Laplacian operator.

If the propagation model does not contain nonlinear effects, steady-state conditions can be assumed, and the Helmholtz equation is used:

$$\nabla^2 u + k^2 u = 0, \tag{2}$$

where $k = \omega/c$ is the wave number, and ω is the angular frequency.

The Fourier transform provides a bridge between the time-domain wave equation and the frequency-domain Helmholtz equation. This relation is used to compare trained PINNs in different domains (Fig. 1). Note that the similarity of solutions in the time and frequency domains is not guaranteed because of using a neural network that is a nonlinear approximator.

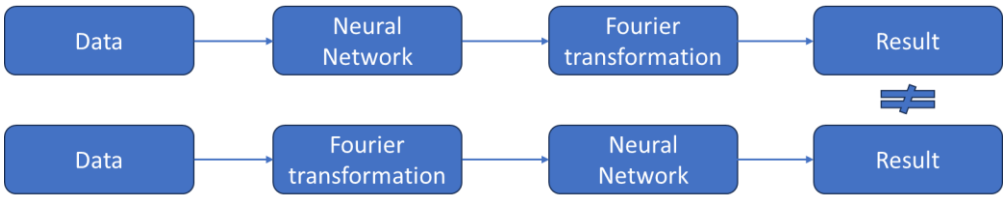


Fig. 1. Comparison of wave propagation in time and frequency domains.

To simplify the investigation, the propagation model of a harmonic wave in a 2D medium is used. The harmonic wave source is located at a fixed point in the domain, and the propagation is computed analytically with a defined spatial and time step. The source is modeled as a harmonic excitation at a fixed point (x_0, y_0) in the domain:

$$u(x, y, t) = \frac{a}{\sqrt{\rho}} \sin(k\rho - \omega t), \tag{3}$$

where $a/\sqrt{\rho}$ is the amplitude of the source wave;
$\rho = \sqrt{(x - x_0)^2 + (y - y_0)^2}$ is the distance from the source of the harmonic wave.
The computed points are used to generate training and test datasets.

The simulation domain is divided into two parts: the core part, which has discretized points from 0 to 1, is used for training and verification of increasing resolution; the area outside of the core is used for verification of extrapolation. The core area has a length of 2.5 wavelength in x and y directions. Data in the core area that is marked by the red rectangle in Fig. 2 has 32 points in the time domain and different discretizations in the space domain, e.g., 32x32 or 8x8. Such greed in the space domain is the worst case for increasing resolution and simple to make conclusions.

The training data are generated for two cases: a single harmonic source that uses the equation above, and two harmonic sources in opposite phase. The goal of the PINN is to predict the wave field across the entire domain, including areas where data are not available.

PINN Architecture and Training

The PINNs used for this study are a fully connected neural network with the following architecture (Fig. 3):

- Input layer: x, y and t for time domain investigation and x, y for frequency domain.
- Hidden layers: three fully connected layers with 64, 128, and 64 neurons, using the Tanh activation function.

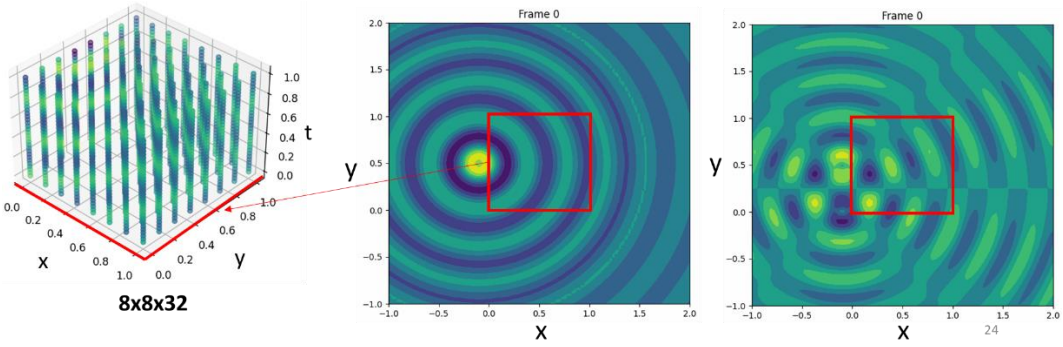


Fig. 2. Input data for neural networks training for single and two sources models.

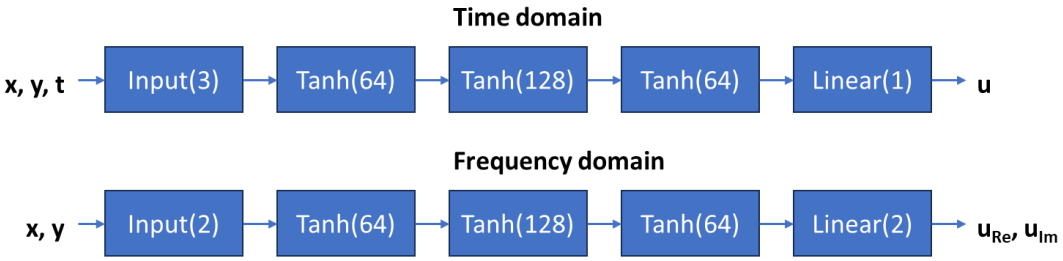


Fig. 3. Architecture of used neural networks for training in time and frequency domains.

- Output layer: one neuron for representing $u(x, y, t)$ in the time domain or two neurons for complex valued $u(x, y)$ in the frequency domain.

The same architecture is used for conventional neural networks with data only losses L_{data}. On the other hand, the PINN loss function used in training combines data loss and physics loss:

$$L = L_{\text{data}} + L_{\text{physics}}, \tag{4}$$

where L_{data} represents the error between the model prediction and the analytical data; L_{physics} represents the error in satisfying the governing PDEs.

The physics loss term is computed by substituting the neural network output into the PDE. Note that boundary conditions and initial conditions are not given separately and are partially present in the analytically computed data.

The ordinary NN and PINN models were developed using PyTorch version 2.5.0 [9].

The training data consists of simulated wave propagation data generated from the wave equation and the Helmholtz equation, which are given above. Physics loss is given in the whole area of interest. Different PINNs are trained in the time or frequency domain to assess their performance in:

- Predicting phase and amplitude at non-computed points. Note that FFT is used for temporal output to represent the result in the frequency domain.
- Extrapolating solutions beyond the dataset domain.

The networks are trained using the Adam optimizer with a learning rate of 0.001. A total of 12,000 training iterations are performed. MSE of the physics loss is scaled by 0.01 relative to MSE of the data loss to provide correct training.

RESULTS AND DISCUSSION

Approximation accuracy for different models

The approximation accuracy is tested for simple NN and PINN in the time and frequency domains. Figure 4 contains the mean square error for different NNs depending on the number of measurements per wavelength. In other words, there is a dependance of error on resolution. It is obvious that decreasing resolution increases error for all cases. However, PINNs approximate data better in the case of decreasing resolution. PINNs stop working if the sampling rate is less than the Nyquist rate (< 2 in Fig. 4).

Note that this result is valid for PINNs in time and frequency domains. So, PINNs can be used to increase resolution, which decreases the number of measurements.

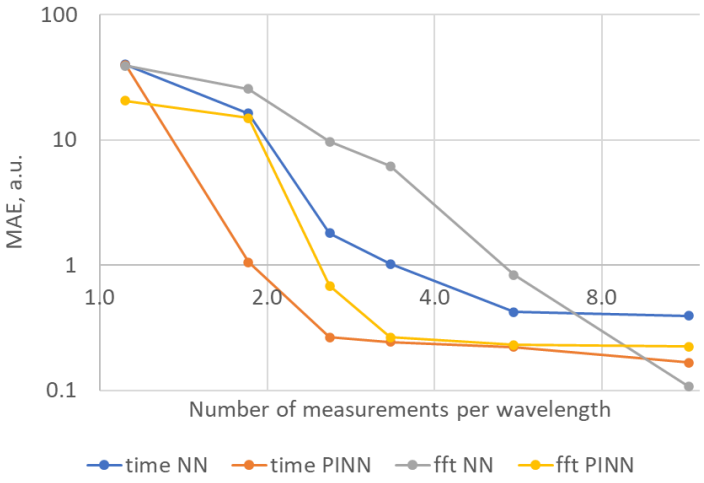


Fig. 4. Approximation error of different NNs depending on data density.

Extrapolation Capability

Ordinary NN and PINN are trained in limited regions and demonstrate different extrapolation performance in the frequency domain. Fig. 5 shows that an ordinary NN gives expectedly incorrect results for the single source.

On the other hand, Fig. 6 shows that PINNs can predict phase and amplitude in neighboring regions where phase is predicted with 20% error up to 0.5 of the training cell size, which is marked by a red rectangle. An exception is the neighboring region that is directed to the source. This is caused by the absence of the wave excitation source in the partial differential equation.

Using PINNs in the frequency domain for predicting narrow band signals allows for a reduction in computational resources.

Extrapolation of PINN in the time domain predicts phase correctly at a distance of 1 cell except the direction to the source (Fig. 7). Amplitude is predicted with an error of < 20% at a distance of 0.5 cells except the direction to the source. Large amplitudes cannot be correctly described by the model in both domains.

Also, Fig. 7 illustrates on the bottom images, predicted wave propagation in the time domain for two sources in opposite phase. It has a similar result to a single source. And these results are better than those for the frequency domain. This allows us to reduce the amount of data for training.

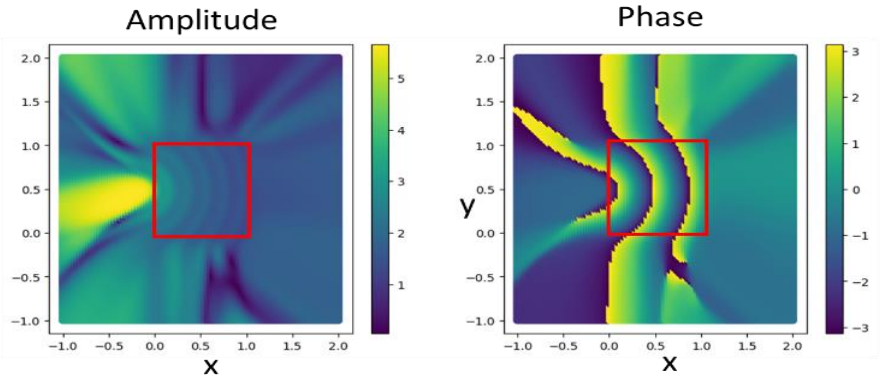


Fig. 5. Extrapolation results in frequency domain of ordinary NN for single source.

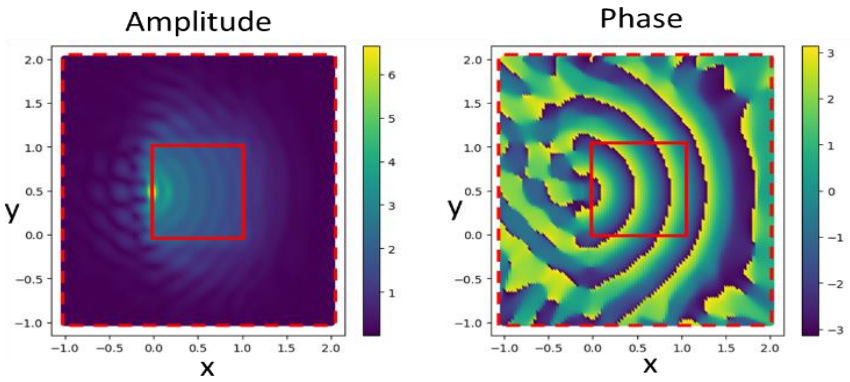


Fig. 6. Extrapolation results in frequency domain of PINN for single source.

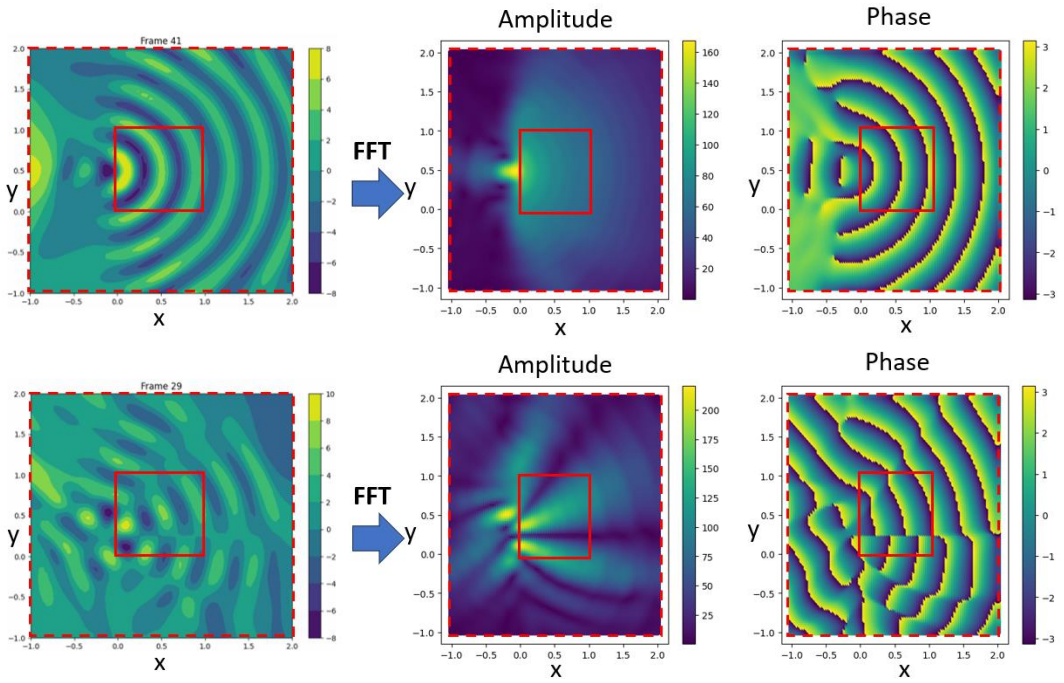


Fig. 7. Extrapolation results in time domain of PINN for single (top) and two (bottom) sources.

Despite the advantages of PINNs, it exhibits the following limitation. Nonlinearities in the neural network architecture prevent exact equivalence between time-domain and frequency-domain solutions.

CONCLUSION

PINNs offer a powerful tool for modeling wave propagation in scenarios where measurement data is limited in some areas. This study demonstrates their ability to improve resolution, extrapolate solutions, and reduce the need for additional data and computations. While challenges remain, particularly in the accurate prediction of large amplitudes and the preservation of solution identity across domains, PINNs represent a promising approach for applications in navigation and wireless tracking.

Future work should focus on improving the accuracy of predictions through better network architectures or hybrid approaches.

REFERENCES

[1] Raissi M., Perdikaris P., & Karniadakis G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686-707. https://doi.org/10.1016/j.jcp.2018.10.045

[2] Chen Y., Lu L., Karniadakis G. E., Negro L. D. (2020). Physics-informed neural networks for inverse problems in nano-optics and metamaterials. Opt. Express 28, 11618-11633 https://doi.org/10.1364/OE.384875

[3] Trahan C., Loveland M. and Dent S. (2024) Quantum Physics-Informed Neural Networks. Entropy, 26(8), 649. https://doi.org/10.3390/e26080649

[4] Gao H., Sun L., Wang J. (2020) Super-resolution and denoising of fluid flow using physics-informed convolutional neural networks without high-resolution labels. arXiv preprint https://arxiv.org/pdf/2011.02364

[5] Moseley B., Markham A., Nissen-Meyer T. (2020) Solving the wave equation with physics-informed deep learning. arXiv preprint https://arxiv.org/pdf/2006.11894

[6] Arzani A., Wang J., D'Souza R. (2021) Uncovering near-wall blood flow from sparse data with physics-informed neural networks. arXiv preprint https://arxiv.org/pdf/2104.08249

[7] Moseley B., Nissen-Meyer T. and Markham A. (2023) Finite Basis Physics-Informed Neural Networks (FBPINNs): a scalable domain decomposition approach for solving differential equations, Advances in Computational Mathematics https://doi.org/10.1007/s10444-023-10065-9

[8] Kolych I., Uln K. (2023) Determination and tracking of trajectories of moving objects in wireless applications - US Patent App. 17/733,919, https://patents.google.com/patent/US20230353980A1/en

[9] PyTorch v2.5.0. Previous PyTorch Versions. URL: https://pytorch.org/get-started/previous-versions/#v250

ФІЗИЧНО-ІНФОРМОВАНІ НЕЙРОННІ МЕРЕЖІ ДЛЯ МОДЕЛЮВАННЯ ПОШИРЕННЯ ВУЗЬКОСМУГОВИХ СИГНАЛІВ

Ігор Колич

Львівський національний університет імені Івана Франка,
вул. Драгоманова, 50, м. Львів, 79005, Україна

АНОТАЦІЯ

Вступ. Фізично-інформовані нейронні мережі (ФІНМ) продемонстрували потужні можливості у вирішенні прямих і обернених задач для диференціальних рівнянь у часткових похідних. У цьому дослідженні основна увага приділяється застосуванню ФІНМ для апроксимації та екстраполяції поширення вузькосмугового сигналу. Ці зусилля мотивовані потенційною можливістю зменшити витрати на вимірювання та чисельні витрати в таких застосунках, як навігаційні системи на основі акустичних і електромагнітних маяків. Ці системи створюють карту середовища та відстежують траєкторії об'єктів, використовуючи дані про поширення хвиль.

Матеріали та методи. Поширення гармонійних хвиль у середовищі можна описати за допомогою хвильового рівняння або рівняння Гельмгольца. Для встановлення зв'язку між цими рівняннями використовується перетворення Фур'є. ФІНМ навчені в часовій або частотній області для прогнозування характеристик поширення хвилі,

таких як амплітуда та фаза. Дослідження порівнює продуктивність ФІНМ зі звичайними нейронними мережами.

Результати. Дослідження показує, що ФІНМ демонструють кращу продуктивність у порівнянні зі звичайними нейронними мережами, коли точки тренувальних даних рознесені до частоти Найквіста. У часовій області ФІНМ точно передбачають фазу на відстані до однієї комірки, за винятком напрямку на джерело. Однак прогнози амплітуди менш точні, з помилками менше 20% на відстані до 0,5 клітини. Для більших амплітуд моделі важко забезпечити надійні прогнози. Навчання ФІНМ у частотній області потребує менше обчислювальних ресурсів, але продуктивність нижча, ніж у часовій області.

Висновки. ФІНМ пропонують багатообіцяючі переваги для моделювання розповсюдження хвиль у вузькосмугових сигналах, особливо в сценаріях, де дані вимірювань розріджені або локальні. Вони можуть збільшити роздільну здатність, зменшити обсяг необхідних даних і оптимізувати обчислювальну ефективність. Незважаючи на їх обмеження, існує різниця рішень між часовою та частотною областями через нелінійну природу нейронних мереж. Майбутня робота може стосуватися точності прогнозів за допомогою кращої архітектури мереж або гібридних підходів.

Ключові слова: Фізично-інформовані нейронні мережі, ФІНМ, хвилі, надвисока роздільна здатність, глибоке навчання, швидке перетворення Фур'є.

Received / Одержано	Revised / Доопрацьовано	Accepted / Прийнято	Published / Опубліковано
09 April, 2025	28 April, 2025	01 May, 2025	26 June, 2025

Electronics and Information Technologies, 2025, 30, 121–136
http://publications.lnu.edu.ua/collections/index.php/electronics/index

UDC: 004.032.26:623.5.051

LANDMINE RECOGNITION USING A KOLMOGOROV–ARNOLD NEURAL NETWORK BASED ON MAGNETIC SENSING DATA

Ivan Peleshchak *, ***Viacheslav Beltiukov***
Lviv Polytechnic National University,
12 Stepan Bandera Str., 79013, Lviv, Ukraine

Peleshchak, I. & Beltiukov, V. (2025). Landmine Recognition Using a Kolmogorov–Arnold Neural Network Based on Magnetic Sensing Data. *Electronics and Information Technologies*, 30, 121–136. https://doi.org/10.30970/eli.30.10

ABSTRACT

Background. Landmine contamination remains a critical issue for more than 60 countries, including Ukraine, where the recovery of agriculture and infrastructure is hampered by hidden explosive devices. The authors propose a passive land-mine recognition approach that combines measurements of magnetic anomalies obtained with an FLC-100 sensor and a Kolmogorov-Arnold Network (KAN), delivering high accuracy while minimising the risk of detonation.

Materials and Methods. A baseline data set of 338 real measurements (sensor voltage V, sensor height H, soil type S) was balanced by generating 50 synthetic records for every “soil–mine” pair using a parameterised normal distribution. After normalising V and applying one-hot encoding to S and the mine classes M, a three-dimensional feature space was formed. Two KAN architectures were evaluated: KAN (3-16-16-4) and KAN (3-64-64-4), both employing cubic B-splines to achieve high-precision mine recognition (> 95 % accuracy). Training was conducted in PyKAN (PyTorch backend) for 65 epochs with a fixed spline grid (k = 3, m = 10) using the Adam optimiser.

Results and Discussion. The KAN (3-16-16-4) model achieved an accuracy of 93.56 % without overfitting; the main confusion occurred between the “anti-personnel” and “booby-trap” classes. Increasing the number of neurons in each hidden layer to 64 raised the accuracy to 95.59 % and eliminated the erroneous assignment of “anti-personnel” mines to the “booby-trap” class. Both networks perfectly distinguished the “no-mine” and “anti-tank” cases, confirming the robustness of spline activations to sensor noise.

Conclusion. The computer experiment shows that a Kolmogorov–Arnold neural network with cubic B-spline weights provides robust recognition of different mine types (“no mine”, “anti-tank mine”, “anti-personnel mine”, “booby-trap”) using magnetic-field sensor data (10^{-10} Tesla) with accuracy exceeding 95 %. Interpretable spline weights allow the contribution of each feature to be analysed, ensuring high sensitivity to small anomalies and demonstrating the scalability of KAN.

Keywords: Kolmogorov–Arnold network, fluxgate magnetic sensor, passive mine detection

INTRODUCTION

Landmine detection remains a persistent and growing global concern, posing life-threatening risks to millions of people. According to the Landmine Monitor 2023, landmines and explosive remnants of war (ERW) continue to cause severe humanitarian consequences, with over 4,700 casualties reported globally in 2022 alone, the vast majority of whom were civilians. More than 60 countries remain contaminated by landmines, presenting

ongoing risks for local populations, especially in post-conflict regions such as Ukraine, where land access, agricultural activity, and reconstruction are critically hindered [1]. In post-war Ukraine, the problem of landmine contamination has become especially urgent, with vast areas of agricultural and residential land requiring safe clearance. Traditional mine detection techniques often lack the reliability and responsiveness needed for large-scale humanitarian demining. Moreover, many active detection methods – based on emitting electric signals – risk triggering explosive devices, endangering human operators.

A promising alternative is the use of passive detection systems [2], particularly those based on magnetic field anomaly sensing [3], [4]. To enhance detection accuracy and reduce operational risks, modern solutions increasingly rely on machine learning techniques, including neural networks. However, neural networks often struggle with noise and distortions in real-world sensor data. One method for improving the robustness and pattern recognition capability of neural architectures is to use neural networks with embedded spline-based functional components, such as Kolmogorov–Arnold Networks (KANs), which are particularly effective at handling noisy and irregular data due to their ability to learn smooth, localised approximations of complex functions.

The aim of this study is to develop an optimised architecture of Kolmogorov-Arnold Networks in terms of the number of hidden layers, neurons per layer, and spline shape for accurate recognition of mine types in soils of varying composition. This research is highly relevant in the context of post-war recovery efforts in Ukraine, where effective and safe detection of minefields plays a crucial role in restoring civil infrastructure and ensuring public safety.

MATERIALS AND METHODS

Literature Review

Kolmogorov–Arnold Networks (KAN) represent a recent advancement in neural network architecture inspired by the Kolmogorov–Arnold representation theorem. Unlike traditional multilayer perceptrons (MLPs) that rely on fixed activation functions at each node, KAN replaces every weight with a univariate, spline-parametrised function. This allows KANs to learn richer functional representations with fewer parameters, making them both efficient and interpretable [5].

In the work by Erdmann et al. [6], KAN was applied to a binary classification problem in high-energy physics. The authors found that while multilayer KANs did not always outperform standard MLPs in terms of accuracy, they demonstrated greater interpretability. Specifically, the activation functions learned in deeper KANs differed significantly from those in shallow models, indicating the architecture's capacity for more abstract feature extraction.

Somvanshi et al. [7] provide a comprehensive survey on KAN, outlining its theoretical foundations and practical adaptations across domains such as biomedical analytics, time series prediction, and graph learning. They highlight KAN’s flexibility and adaptability, particularly in handling high-dimensional structured data.

Barasin et al. [8] explored KAN in the context of time series classification using the UCR benchmark dataset. Their findings revealed that well-optimised KAN models outperformed MLPs and achieved competitive results compared to state-of-the-art models such as HIVE-COTE2, all while maintaining computational efficiency and robustness to hyperparameter changes.

In terms of robustness, the study published in Applied Sciences assessed the vulnerability of different KAN architectures to adversarial attacks [9]. Among the variants, KAN-Mixer showed the best performance in resisting attacks while retaining strong accuracy on clean data. This makes KAN suitable for safety-critical applications like mine detection, where robustness is paramount.

In the field of remote sensing, Cheon [10] proposed combining pretrained CNNs with KAN layers for scene classification using the EuroSAT dataset. The hybrid models

achieved high classification accuracy with reduced parameter counts and faster training, illustrating the potential for integrating KAN into real-time systems.

Drokin [11] extended KAN's application to computer vision tasks, proposing parameter-efficient KAN convolution layers and fine-tuning techniques. The results demonstrated that KAN-based models can achieve strong performance in both image classification and segmentation tasks, suggesting relevance to image-based mine detection scenarios.

The reviewed literature suggests that KAN offers a unique combination of interpretability, efficiency, and reliability across various classification domains. These characteristics make it a promising candidate for mine detection, especially in post-war Ukraine, where safety, dependability, and explainability are of paramount importance. However, the optimisation of the Kolmogorov-Arnold Network architecture to improve recognition accuracy — as well as the trade-off between training speed and recognition precision —remains an open challenge, which is crucial in the context of mine detection.

Classification data

In this study, we utilised the dataset provided in [4], which focuses on the classification of landmines based on magnetic field anomaly characteristics. The parameter values employed in our experimental setup are summarised in Table 1. Furthermore, we analysed the relationship between magnetic anomaly values and the soil type (Table 2), as well as the distance between the magnetic sensor and the buried landmine (Table 3). The general trends in magnetic field anomalies across different landmine types were also examined and illustrated (Table 3).

To obtain reliable measurements of the magnetic anomalies surrounding subsurface mines, the original study [4] employed a fluxgate magnetic sensor model FLC100 (10^{-10} Tesla) [12], which demonstrated sensitivity to minute variations in the magnetic field. This sensor-based approach enabled passive mine detection without the need for active signal emission, thus reducing the risk of accidental detonation. The design and deployment of the sensing mechanism were previously validated in [4], where a decision support system for mine classification was developed using metaheuristic classifiers.

Table 1. **Parameters data [4]**

<table>
<tr><th colspan="5">The Parameters</th></tr>
<tr><th rowspan="2"></th><th colspan="3">Input Data, “Independent Variables”</th><th>Output Data, “Dependent Variable”</th></tr>
<tr><th>Voltage (V)</th><th>High (H)</th><th>Soil Type (S)</th><th>Mine Type (M)</th></tr>
<tr><td>Definition</td><td>The value of the output voltage of the FLC sensor is due to the action of the magnetic anomaly.</td><td>The distance of the sensor above the ground.</td><td>6 different types of soil depending on the state of moisture.</td><td>Types of mines commonly found on land; 4 different classes of mines.</td></tr>
<tr><td rowspan="6">Limit values/ Class</td><td rowspan="6">[0 V, 10.6 V]</td><td rowspan="6">[0 cm, 20 cm]</td><td>Dry and sandy</td><td>No mine</td></tr>
<tr><td>Dry and purulent</td><td>Anti-tank</td></tr>
<tr><td>Dry and chalky</td><td>Anti-personnel</td></tr>
<tr><td>Wet and sandy</td><td rowspan="3">Booby trapped Anti-personnel</td></tr>
<tr><td>Humid and humus</td></tr>
<tr><td>Wet and chalky</td></tr>
</table>

Table 2. **Dependence of magnetic field anomalies in the vicinity of mines on soil type [4]**

Soil Type	Null, V	Anti-Tank, V	Anti-Personnel, V	Booby Trapped Anti-Personnel, V
Dry and sandy	3.560	10.400	3.830	5.590
Dry and purulent	3.500	7.500	3.920	5.590
Dry and chalky	3.720	10.400	6.890	2.406
Wet and sandy	3.780	10.400	6.220	4.490
Humid and humus	3.350	10.400	5.050	2.770
Wet and chalky	3.610	10.400	5.960	4.400

Preprocessing data

Data generation based on parameterised normal distribution

To improve the generalisation capability of the model on a limited dataset consisting of 338 real records [4], an additional data generation procedure was applied using a parameterised normal distribution.

The chosen method is based on generating new examples by adding pseudorandom noise [13] to the original feature values V and H within each subgroup of data defined by a unique pair of soil type S and mine type M. For each such subgroup, the statistical characteristics of the features are computed as follows:

$$\mu_V = \overline{V}, \qquad \sigma_V = \mathrm{std}(V), \qquad \mu_H = \overline{H}, \qquad \sigma_H = \mathrm{std}(H), \tag{1}$$

where μ_V, μ_H — mean values of features V and H, respectively,
σ_V, σ_H — standard deviations of features V and H,

Table 3. **Dependence of the magnetic field anomaly in the vicinity of mines on the distance from the sensor to the ground [4]**

Height (cm)	Mine Type 1 Voltage, V	Mine Type 2 Voltage, V	Mine Type 3 Voltage, V	Mine Type 4 Voltage, V
0.00	3.6	10.4	4.1	5.9
1.82	3.4	10.4	4.0	5.5
3.64	3.4	10.4	3.8	5.0
5.45	2.8	10.4	3.9	4.4
7.27	2.9	9.5	3.6	4.3
9.09	2.7	8.3	3.4	4.25
10.91	2.9	7.0	3.4	4.2
12.73	2.6	6.4	3.45	4.05
14.55	2.5	6.2	3.5	3.9
16.36	2.6	4.8	3.8	3.2
18.18	2.6	4.6	3.2	3.2
20.00	2.4	4.5	3.2	3.1

V, H — arithmetic means of the respective columns,
$\mathrm{std}(X)$ — standard deviations of features V and H.

New examples are generated using the following formulas:

$$V_i' = V_i + N(0, \sigma_V \cdot \alpha),$$
$$H_i' = H_i + N(0, \sigma_H \cdot \alpha), \quad (2)$$

where V_i', H_i' — newly generated values of magnetic field anomaly and height for the i-th sample,
V_i, H_i — values sampled from an existing record in the subgroup,
$\alpha = 0.1$ — noise intensity coefficient (empirically selected),
$N(0, \sigma)$ — normally distributed random value with mean 0 and standard deviation σ.

To ensure the physical plausibility of generated values, clipping was applied to constrain them within real-world sensor bounds:

$$V_i' \in [0.0V, 10.6V], \qquad H_i' \in [0.0cm, 20.0cm], \quad (3)$$

in accordance with the sensor specifications.

For each subgroup defined by (S, M), 50 new samples were generated, which significantly increased the number of training examples and smoothed the data distribution.

Data preprocessing before training

Before training the neural network, the following preprocessing steps were performed:

1. Normalisation of magnetic anomaly feature V:

$$V' = \frac{V - \overline{V}}{\sigma_V}, \quad (4)$$

where V'— normalised magnetic anomaly value,
V — original value of the magnetic anomaly,
$\overline{V}$ — mean magnetic anomaly over the entire dataset,
σ_V — standard deviation of the magnetic anomaly.

2. Categorical encoding of the soil type variable S, which takes six values:
   - “Dry and Sandy”.
   - “Dry and Humus”.
   - “Dry and Limy”.
   - “Humid and Sandy”.
   - “Humid and Humus”.
   - “Humid and Limy”.

These categories were encoded using **One-Hot Encoding**, which transforms each category into a binary vector of size six. For example, if the soil type is the second category (“Dry and Humus”), the vector would be:

$$[0,1,0,0,0,0]. \quad (5)$$

3. Target encoding. The mine types M, originally in categorical form, were first mapped to numerical indices (0–4) and then encoded using one-hot encoding for input into the neural network.

This method preserved the internal structure and semantics of the data, ensured physical interpretability of the generated values, and significantly improved the model's generalisation potential.

Mathematical Model of Kolmogorov–Arnold Networks

The articles [3], [4] address a multiclass classification task solved with classical machine-learning techniques — artificial neural networks [14] and their variants, support-vector machines [15], Bayesian approaches [16], decision trees, and others [17].

Let X be the feature space X = {V, H, S}, where V denotes the magnetic-field anomaly in the vicinity of a mine (volts), H is the sensor height above the ground that covers the mine, and S represents the soil type. The label set is Y = {0, 1, 2, 3}, whose elements correspond to the classes "no mine," "anti-tank mine," "anti-personnel mine," and "booby-trap," respectively.

The classification objective is to determine a mapping operator Y*: X → Y that assigns any previously unseen object x ∈ X to class y ∈ Y while minimising the Euclidean error

$$min \parallel y^* - y \parallel, \tag{6}$$

where y the true class label and y^* is the neural-network prediction [14].

KAN morphology

The Kolmogorov–Arnold Network (KAN) is a neural architecture inspired by the Kolmogorov-Arnold representation theorem [18]. Unlike traditional MLPs that apply fixed nonlinearities at nodes and learn linear weights, KANs apply **learnable nonlinear activation functions on edges**, modelled as univariate splines. Each layer in a KAN consists of **a matrix of spline functions**, and each neuron simply sums the outputs of these spline-parameterised edges.

General architecture

We consider KAN architecture with the shape [3, 16, 16, 4] (Fig. 1), which includes an input layer comprising 3 nodes corresponding to the input features, followed by two hidden layers with 16 nodes each, and an output layer consisting of 4 nodes representing either classification labels or regression targets.

The general forward propagation is expressed by the composition of KAN layers:

$$KAN(x) = (\Phi_2 \circ \Phi_1 \circ \Phi_0)(x), \tag{7}$$

where each Φ_l is a **functional matrix** consisting of learnable spline activations.

Every layer transforms its input by applying these univariate spline functions on each edge, followed by summation at the next layer's nodes.

Layer-wise formulation

Let the input vector be $x^{(0)} = x \in \mathbb{R}^3$. The subsequent layer computations are defined as follows:

1. **First Hidden Layer (Layer 0 → 1).**

For each neuron $j = 1,\ldots,16$ in the first hidden layer:

$$x_j^{(1)} = \sum_{i=1}^{3} \varphi_{j,i}^{(0)}\left(x_i^{(0)}\right). \tag{8}$$

2. **Second Hidden Layer (Layer 1 → 2).**

For each neuron $k = 1,\ldots,16$ in the second hidden layer:

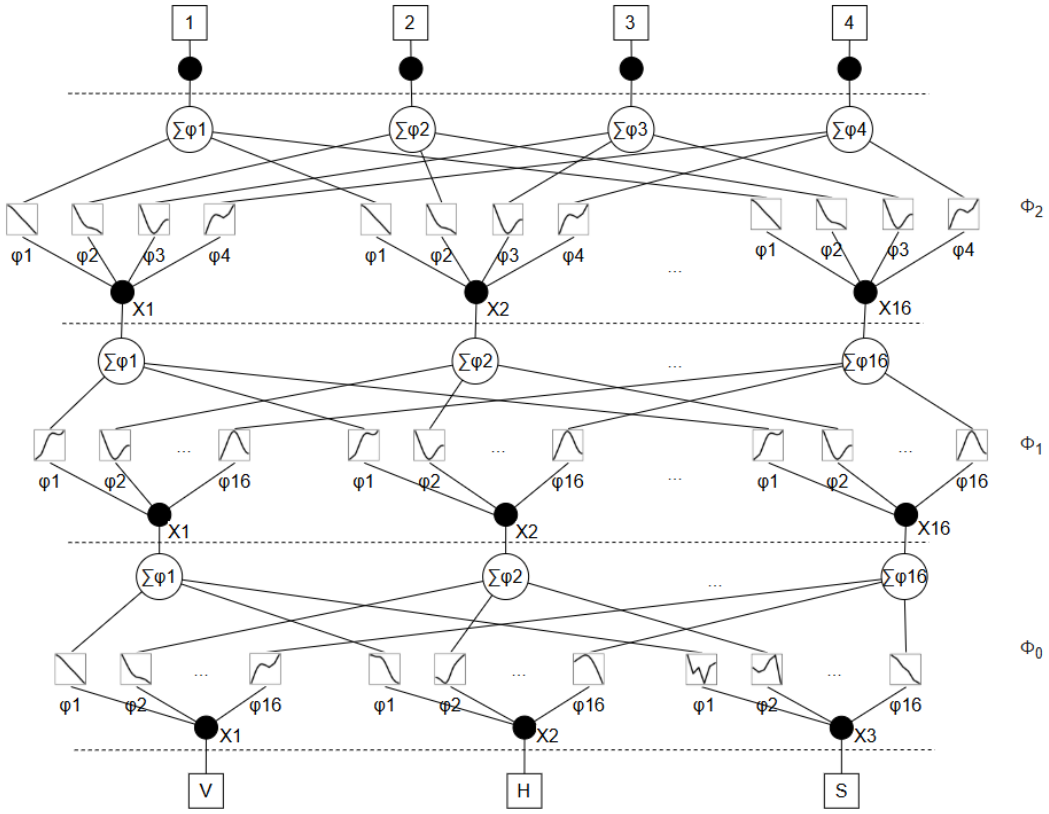


Fig. 1. General architecture of KAN.

$$x_k^{(2)} = \sum_{j=1}^{16} \varphi_{k,j}^{(1)}\left(x_j^{(1)}\right). \tag{9}$$

3. Output Layer (Layer 2 → 3).

For each output node $m = 1,\ldots,4$:

$$x_k^{(2)} = \sum_{j=1}^{16} \varphi_{k,j}^{(1)}\left(x_j^{(1)}\right). \tag{10}$$

The final output of the model is:

$$KAN(x) = \begin{bmatrix} x_1^{(3)} \\ x_2^{(3)} \\ x_3^{(3)} \\ x_4^{(3)} \end{bmatrix} \in \mathbb{R}. \tag{11}$$

Spline Activation Functions

Each edge activation function $\varphi_{j,i}^{(l)}(x)$ is defined as a combination of a residual nonlinear term and a cubic B-spline [19]:

$$\varphi(x) = w_b \cdot \mathrm{silu}(x) + w_s \cdot \sum_{m=0}^{G+k-1} c_m B_m(x). \quad (12)$$

where $\mathrm{silu}(x) = x/(1+e^{-x})$ is the smooth SiLU function (acts as a residual base),
$B_m(x)$ are cubic B-spline basis functions (order $k = 3$),
$G = 10$ is the number of intervals $\rightarrow G + k = 13$ of basis functions per spline,
c_m are trainable spline coefficients,
w_b, w_s are trainable scalar weights controlling the contribution of the SiLU and the spline.

Spline Activation Functions

The total number of parameters can be calculated as follows. Each spline contains $G + k = 13$ coefficients, in addition to two weights (w_b, w_s), resulting in 15 parameters per edge. The first layer (φ_0) contains $3 \times 16 = 48$ edges, contributing $48 \cdot 15 = 720$ parameters. The second layer (φ_1) has $16 \times 16 = 256$ edges, yielding $256 \cdot 15 = 3840$ parameters. The third layer (φ_2) includes $16 \times 4 = 64$ edges, resulting in $64 \cdot 15 = 960$ parameters. Therefore, the total number of parameters is $720 + 3840 + 960 = 5520$.

Computational Environment and Tools

The details of the experimental environment, including the software tools and libraries, are as follows:

Programming language: Python

Neural network library: PyKAN [20]

Hardware: Personal PC (AMD Ryzen 5 5600G CPU, NVIDIA GeForce RTX 4060 GPU)

Software:

- IntelliJ IDEA (with Python plugin support).
- Python 3.x.
- PyKAN library [20].
- CUDA Toolkit (for GPU acceleration with NVIDIA RTX 4060).
- PyTorch (backend library for PyKAN).
- NumPy (for data manipulation).
- Matplotlib (for visualisation of results).

RESULTS AND DISCUSSION

In this section, we present the setup and execution of the computer experiment aimed at evaluating the performance of Kolmogorov–Arnold Networks (KANs) for a multiclass classification task. The experiment involved training and comparing two network architectures: KAN (3,16,16,4) and KAN (3,64,64,4). The architectures were chosen based on the task's dimensionality, where the input space was three-dimensional, and the output space consisted of four distinct classes.

The training was conducted using the PyKAN library [20] on the Python platform. The settings for the networks included the use of cubic B-splines as basis functions, with the order set to 3 and the grid size set to 10. These parameters provided a sufficient balance between the flexibility of the spline approximation and computational efficiency.

The network KAN (3,16,16,4) was trained first. It underwent a training process over a 65 epochs and achieved an accuracy of 93.56% on the test set. In the second case, the KAN (3,64,64,4) architecture was trained. Due to its significantly higher number of parameters, the training took a considerably longer time; however, it achieved an improved test accuracy of approximately 95,59%.

During the training process, loss and accuracy curves were recorded for each model to monitor convergence dynamics and to detect potential signs of overfitting. After the evaluation phase, confusion matrices were generated to provide a detailed understanding of classification performance across all classes. In addition to the visual analyses, a comprehensive classification report was produced, presenting key metrics such as precision, recall, and F1-score for each class.

The results of the experiments are illustrated through loss-accuracy curves, confusion matrices, and a set of other performance metrics [21], which comprehensively describe the behavior of both tested architectures.

For the KAN (3,16,16,4) network, the loss curve (Fig. 2) demonstrated a steady decrease without abrupt oscillations, indicating stable convergence. The corresponding accuracy curve (Fig. 2) showed consistent improvement throughout the training process, reaching a plateau near 92.56%. The confusion matrices (Fig. 3) revealed that most misclassifications occurred between the (specify which classes if possible), suggesting that the network found these classes harder to differentiate given the feature space.

The normalized confusion matrix for the KAN (3–16–16–4) model, shown in Fig. 3b, reflects almost perfect identification of the "no mine" and "anti-tank" classes, with correct detection rates reaching approximately ninety-seven percent. At the same time, the majority of misclassifications occurred between the "anti-personnel" and "booby trap" classes: around ten objects from the first category were confused with the second, while the reverse misclassification happened almost twice as rarely. This asymmetry is explained by the partial overlap of magnetic anomaly ranges and sensor height, indicating that the three-dimensional feature space was insufficient to fully separate these mine types.

Despite this, the model exhibits stable convergence of the loss function and absence of sharp fluctuations, indicating proper hyperparameter tuning and sufficient capacity for the basic task. However, it also signals the need to enrich the feature space specifically in the area where classification errors are observed.

On the other hand, the KAN (3,64,64,4) architecture, while requiring longer training time due to the increased number of neurons, achieved superior classification results with approximately 95% accuracy. Its loss curve (Fig. 4) exhibited a smoother descent, and its accuracy curve (Fig. 4) achieved a slightly higher and more stable plateau compared to

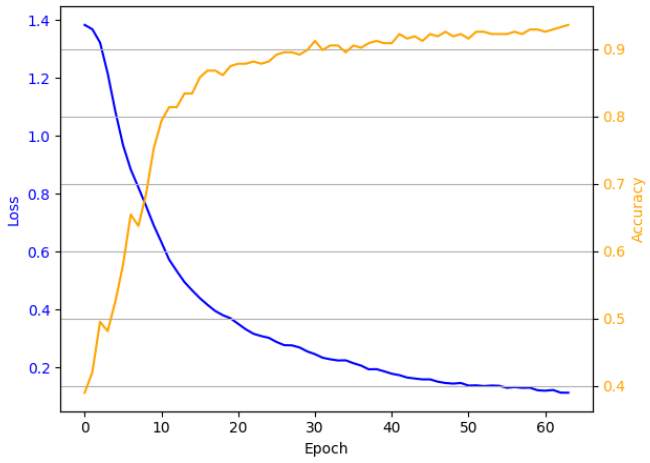


Fig. 2. Training accuracy and loss over epochs for KAN (3, 16, 16, 4).

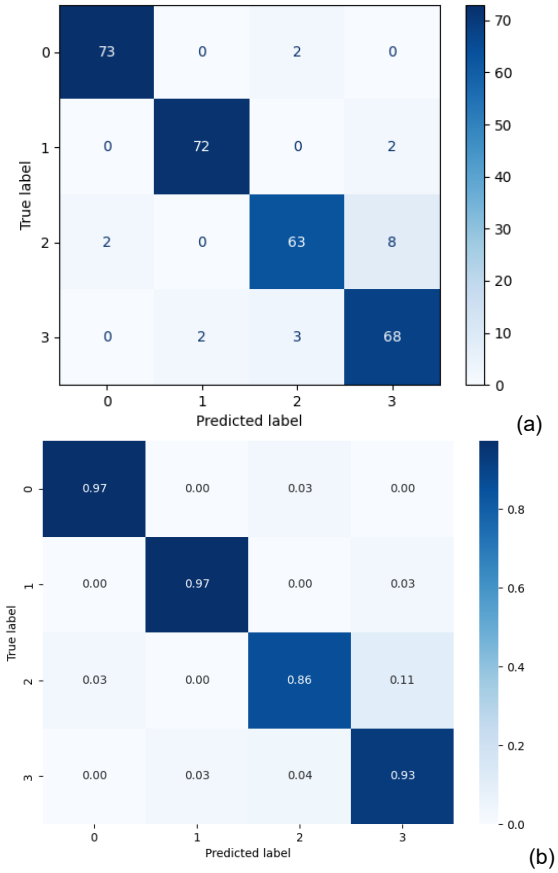


Fig. 3. Confusion matrix (a) and normalized confusion matrix (b) for KAN (3, 16, 16, 4).

the smaller network. The confusion matrices (Fig. 5) for this model showed a significant reduction in misclassification rates across all classes, particularly improving recognition of “anti-personnel mine” class.

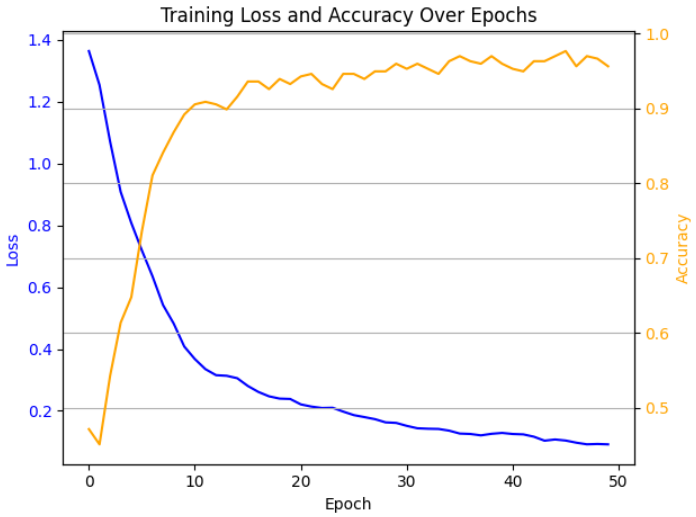


Fig. 4. Loss and accuracy over epochs for KAN (3, 64, 64, 4).

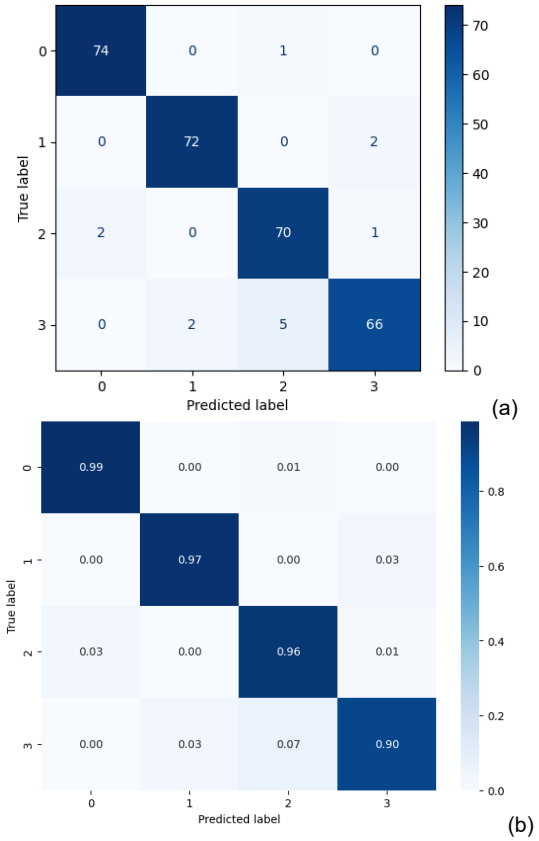


Fig. 5. Confusion matrix (a) and Normalized confusion matrix (b) for KAN (3, 64, 64, 4).

Increasing the number of neurons to sixty-four in each hidden layer led to significant changes in the error patterns, as clearly seen in the confusion matrices in Fig. 5. The updated KAN (3–64–64–4) architecture almost completely eliminated confusion between the "anti-personnel" and "booby trap" classes in the direction from the latter to the former, raising the accuracy for the "booby trap" class above ninety-five percent. Reverse confusion still occurred in about seven cases out of seventy-three, reducing the recall of this class to ninety-four percent, but these mistakes now have a one-sided nature: the network becomes more conservative, assigning doubtful samples to the less dangerous category in the absence of convincing evidence. The increased computational costs are justified by the fact that overall classification accuracy improved by about two percent, and the off-diagonal elements of the matrix sharply decreased for all classes except for the localised issue of booby trap identification.

The comparison of the two models shows that even with the same basic set of features, a wider architecture can capture finer signal nonlinearities and thus reduce the number of critical errors. At the same time, the remaining confusion between classes 2 and 3 indicates the limit beyond which pure network scaling becomes less effective compared to introducing additional information, such as gradient characteristics of the magnetic field or contextual soil indicators.

Thus, detailed analysis of the confusion matrices indicates that the main direction for further optimization should be strengthening the discriminative power of features specifically for the "booby trap" class, while preserving the already achieved high reliability in detecting other mines and safe areas.

Comparative analysis of the two models (Table 4 and Table 5) indicates that increasing the hidden layer size improves generalization capability but at the cost of greater computational time and resources. This trade-off must be considered depending on the application domain requirements.

Table 4. **Classification Report for KAN (3, 16, 16, 4)**

Class	Precision	Recall	F1-score	Support
0	0.9733	0.9733	0.9733	75
1	0.9730	0.9730	0.9730	74
2	0.9265	0.8630	0.8936	73
3	0.8718	0.9315	0.9007	73
Accuracy	–	–	0.9356	295
Macro average	0.9361	0.9352	0.9351	295
Weighted average	0.9365	0.9356	0.9355	295

Notes:

Columns:

- Precision – The proportion of predicted positive samples that are actually correct for each class.
- Recall – The proportion of actual positive samples that are correctly predicted for each class.
- F1-score – The harmonic mean of precision and recall for each class, providing a balance between the two metrics.
- Support – The number of true instances for each class in the test set.

Rows:

- 0, 1, 2, 3 – The performance metrics for the classes “no mine,” “anti-tank mine,” “anti-personnel mine,” and “booby-trap,” respectively.
- Accuracy – The overall classification accuracy across all classes (i.e., the proportion of correctly classified samples).
- Macro average – The unweighted mean of precision, recall, and F1-score across all classes, treating each class equally regardless of its support.
- Weighted average – The mean of precision, recall, and F1-score weighted by the number of true instances (support) for each class, giving more influence to classes with more samples.

Table 5. **Classification report for KAN (3, 64, 64, 4)**

Class	Precision	Recall	F1-score	Support
0	0.9737	0.9867	0.9801	75
1	0.9730	0.9730	0.9730	74
2	0.9211	0.9589	0.9396	73
3	0.9565	0.9041	0.9296	73
Accuracy	–	–	0.9559	295
Macro average	0.9561	0.9557	0.9556	295
Weighted average	0.9562	0.9559	0.9558	295

Thus, the computer experiments have demonstrated that Kolmogorov–Arnold Networks, when properly configured with cubic B-splines and an appropriate grid resolution, can achieve high accuracy in multiclass classification problems, with performance scaling positively with network capacity.

Study Limitations

The base dataset has a limited volume; although synthetic augmentation improves generalization, it cannot fully replace field measurements. The results were obtained under laboratory conditions without considering the influence of metallic debris, heterogeneous magnetic backgrounds, or sensor temperature drifts.

Future Research Directions

Collection of large-scale field data under various climatic and geological conditions to validate the results.

End-to-end optimization: selection of spline grids, nonlinearity bases, and regularization techniques (e.g., KAN-Mixer, weight priorities) to further improve accuracy without exponential growth in parameters.

Robustness: investigation of resilience to adversarial influences typical of deceptive mine masking with metallic shrapnel or geomagnetic traps.

CONCLUSION

1. For the first time, Kolmogorov–Arnold Networks (KAN) with cubic B-splines were used for passive mine recognition based on magnetic anomalies. Unlike classical MLPs, KAN allows modelling nonlinear dependencies at the level of weight connections, enhancing the interpretability and robustness of the model to noise in real sensor measurements.

2. An extended dataset was created: synthetic samples were added to 338 original magnetic anomaly recordings, generated using a parameterized normal distribution with a step of 50 samples for each "soil type – mine type" subset. This balanced the feature variance and reduced the risk of overfitting.

3. Two architectures were developed: KAN (3 - 16 - 16 - 4) and KAN (3 - 64 - 64 - 4). Both models were trained using the PyTorch-compatible PyKAN library with identical spline hyperparameters.

KAN (3 - 16 - 16 - 4) achieved 93.56% accuracy without signs of overfitting; the main errors occurred between the "anti-personnel" and "booby trap" classes. Increasing the number of neurons in KAN (3 - 64 - 64 - 4) to 64 per hidden layer improved accuracy to 95.59%, significantly reducing false detections across all four classes. The cost of this improvement was an almost linear increase in the number of parameters and training time. This confirms the advisability of adaptively selecting model size based on the hardware constraints of field systems. Confusion-matrix analysis showed both models nearly flawless at identifying “no-mine” and “anti-tank” cases, while most errors arose from confusion between anti-personnel mines and booby-traps. Results confirm that increasing network capacity improves discrimination among visually similar magnetic signatures.

4. Advantages of the proposed approach.

Passive mine recognition: the use of the FLC-100 sensor does not require active excitation, minimizing the risk of detonation.

Interpretability: spline weights enable analysis of the contribution of each feature and facilitate safety certification.

Robustness: experiments showed no sharp fluctuations in the loss function and stable convergence even on a noise-enriched dataset.

The study proves that properly configured Kolmogorov–Arnold Networks can achieve over 95% accuracy in multiclass mine classification based on passive magnetic features. The combination of interpretable spline weights, high sensitivity to small anomalies, and scalability potential makes KAN a promising foundation for modular humanitarian demining systems, which can significantly accelerate land clearance and reduce risks for personnel.

ACKNOWLEDGMENTS AND FUNDING SOURCES

The research was carried out with the grant support of the National Research Fund of Ukraine "Methods and means of active and passive recognition of mines based on deep neural networks", project registration number 273/0024 from 1/08/2024 (2023.04/0024). Also, we would like to thank the reviewers for their precise and concise recommendations that improved the presentation of the results obtained.

COMPLIANCE WITH ETHICAL STANDARDS

Conflict of Interest: The authors declare that the research was conducted in the absence of any.

AUTHOR CONTRIBUTIONS

Conceptualization, [I.P.]; methodology, [I.P.]; validation, [I.P.]; formal analysis, [I.P., V.B.].; investigation, [V.B.]; resources, [V.B.]; data curation, [V.B.]; writing – original draft preparation, [V.B.]; writing – review and editing, [I.P.]; visualization, [V.B.]; supervision, [I.P.]; project administration, [I.P.]; funding acquisition, [I.P.].

REFERENCES

[1] Landmine Monitor 2023. (2023). https://backend.icblcmc.org/assets/reports/Landmine-Monitors/LMM2023/Downloads/Landmine-Monitor-2023_web.pdf

[2] Lytvyn, V., Peleshchak, I., Peleshchak, R., Mediakov, O., & Pukach, P. (2023). Development of a hybrid neural network model for mine detection by using ultrawideband radar data. Eastern-European Journal of Enterprise Technologies, 3(9 (123)), 78–85. https://doi.org/10.15587/1729-4061.2023.279891

[3] Peleshchak, R. M., Lytvyn, V. V., Nazarkevych, M. A., Peleshchak, I. R., & Nazarkevych, H. Y. (2024). Influence of the symmetry neural network morphology on the mine detection metric. Symmetry, 16(4), 485. https://doi.org/10.3390/sym16040485

[4] Yilmaz, C., Kahraman, H. T., & Soyler, S. (2018). Passive mine detection and Classification Method based on hybrid model. IEEE Access, 6, 47870–47888. https://doi.org/10.1109/access.2018.2866538

[5] Liu, Z., Wang, Y., Vaidya, S., Ruehle, F., Halverson, J., Soljačić, M., Hou, T. Y., & Tegmark, M. (2024). KAN: Kolmogorov-Arnold Networks. arXiv. https://doi.org/10.48550/arxiv.2404.19756

[6] Erdmann, J., Mausolf, F., & Späh, J. L. (2024). KAN we improve on HEP classification tasks? Kolmogorov-Arnold Networks applied to an LHC physics example. arXiv. https://doi.org/10.48550/arxiv.2408.02743

[7] Somvanshi, S., Javed, S. A., Islam, M. M., Pandit, D., & Das, S. (2024). A survey on Kolmogorov-Arnold Network. arXiv. https://doi.org/10.48550/arxiv.2411.06078

[8] Barašin, I., Bertalanič, B., Mohorčič, M., & Fortuna, C. (2024). Exploring Kolmogorov-Arnold networks for Interpretable Time Series Classification. arXiv. https://doi.org/10.48550/arxiv.2411.14904

[9] Wang, S., Yue, X., Lyu, Y., & Shan, C. (2025). Exploring Adversarial Transferability between Kolmogorov-arnold Networks. arXiv.org. https://doi.org/10.48550/arXiv.2503.06276

[10] Cheon, M. (2024). Kolmogorov-Arnold Network for Satellite image classification in remote sensing. arXiv. https://doi.org/10.48550/arxiv.2406.00600

[11] Drokin, I. (2024). Kolmogorov-Arnold Convolutions: Design Principles and Empirical Studies. arXiv. https://doi.org/10.48550/arxiv.2407.01092

[12] Behzad. (2018). Data Sheet FLC 100. Scribd. https://scribd.com/document/382854526/Data-Sheet-FLC-100

[13] Maharana, K., Mondal, S., & Nemade, B. (2022). A review: Data pre-processing and data augmentation techniques. Global Transitions Proceedings, 3(1), 91–99. https://doi.org/10.1016/j.gltp.2022.04.020
[14] Qamar, R., & Zardari, B. A. (2023). Artificial Neural Networks: An Overview. Mesopotamian Journal of Computer Science, 130–139. https://doi.org/10.58496/mjcsc/2023/015
[15] Vanneschi, L., & Silva, S. (2023). Support vector machines. In Natural computing series, 271–281. https://doi.org/10.1007/978-3-031-17922-8_10
[16] Lampinen, J., & Vehtari, A. (2001). Bayesian approach for neural networks—review and case studies. Neural Networks, 14(3), 257–274. https://doi.org/10.1016/s0893-6080(00)00098-8
[17] De Ville, B. (2013). Decision trees. Wiley Interdisciplinary Reviews Computational Statistics, 5(6), 448–455. https://doi.org/10.1002/wics.1278
[18] Schmidt-Hieber, J. (2020). The Kolmogorov-Arnold representation theorem revisited. arXiv. https://doi.org/10.48550/arxiv.2007.15884
[19] Chaudhuri, A. (2021). B-Splines. arXiv. https://doi.org/10.48550/arxiv.2108.06617
[20] KindXiaoming. (2024). GitHub - KindXiaoming/pykan: Kolmogorov Arnold Networks. GitHub. https://github.com/KindXiaoming/pykan
[21] Eberhart, R. C., Dobbins, R. W., & Hutton, L. V. (1990). Performance Metrics. In Elsevier eBooks, 161–176. https://doi.org/10.1016/b978-0-12-228640-7.50013-1

РОЗПІЗНАВАННЯ МІН ЗА ДОПОМОГОЮ НЕЙРОННОЇ МЕРЕЖІ КОЛМОГОРОВА-АРНОЛЬДА НА ОСНОВІ ДАНИХ МАГНІТНОГО ЗОНДУВАННЯ

Іван Пелещак *, ***В'ячеслав Бельтюков***
Національний університет «Львівська політехніка»,
вул. Степана Бандери, 12, 79013, Львів, Україна

АНОТАЦІЯ

Вступ. Мінна небезпека залишається критичною для понад 60 країн, зокрема України, де відновлення сільського господарства й інфраструктури стримується прихованими вибуховими пристроями. Автори пропонують пасивний підхід до розпізнавання мін, який поєднує вимірювання магнітних аномалій сенсором FLC-100 із нейронною мережею Колмогорова-Арнольда (НМКА), що забезпечує високу точність за мінімального ризику детонації.

Матеріали та методи. Базовий набір даних із 338 реальних вимірювань (напруга V, висота сенсора H, тип ґрунту S) було збалансовано шляхом генерації 50 синтетичних записів для кожної пари "ґрунт-міна" за параметризованим нормальним розподілом. Після нормалізації V та one-hot кодування S і класів мін M сформовано тривимірний простір ознак. Застосовувалися дві архітектури НМКА: НМКА (3-16-16-4) та НМКА (3-64-64-4) з кубічними B-сплайнами для високоточного розпізнавання мін (>95 % точності). Навчання виконувалось за допомогою бібліотеки PyKAN (PyTorch backend) протягом 65 епох із фіксованою сіткою сплайнів (k = 3, m = 10) з використанням оптимізатора Адам.

Результати. Модель НМКА (3-16-16-4) досягла точності 93,56% без перенавчання; головна плутанина спостерігалась між класами «протипіхотна міна» та «міна-пастка». Збільшення кількості нейронів у кожному прихованому шарі до 64 збільшило точність до 95,59 % і усунуло хибні віднесення «протипіхотна міна» до «міна-пастка». Обидві мережі бездоганно розрізняли випадки «без міни» та «протитанкова міна», що підтверджує стійкість сплайнових активацій до шуму сенсора.

Висновки. Комп'ютерний експеримент показав, що нейронна мережа Колмогорова-Арнольда з кубічними B-сплайновими вагами має високу надійність розпізнавання різних типів мін («без міни», «протитанкова міна», «протипіхотна міна», «міна-пастка») з використанням даних сенсорів магнітного поля (10^{-10} – 10^{-4} Тл) з точністю понад 95%.

Показано, що інтерпретовані ваги сплайнів дозволяють аналізувати внесок кожної ознаки, що забезпечує високу чутливість до малих аномалій та масштабованість НМКА.

Ключові слова: мережа Колмогорова-Арнольда, феррозондовий магнітний сенсор, пасивне виявлення мін

Received / Одержано	Revised / Доопрацьовано	Accepted / Прийнято	Published / Опубліковано
04 May, 2025	03 June, 2025	06 June, 2025	26 June, 2025

Electronics and information technologies, 2025, 30, 137–154
http://publications.lnu.edu.ua/collections/index.php/electronics/index

UDC 004.94

PROGRAMS FOR CALCULATING A DYNAMIC MODE MAP ON THE EXAMPLE OF A SYSTEM WITH CHAOTIC MODES

Sergiy Sveleba*, *Ivan Katerynchuk*, *Ivan Kunyo*, *Yaroslav Shmyhelskyy
Ivan Franko National University of Lviv,
107 Tarnavsky St., UA–79017 Lviv, Ukraine

Sveleba S., Katerynchuk, I., Kunyo, I., & Shmyhelskyy, Y. (2025). Programs for Calculating a Dynamic Mode Map on the Example of a System with Chaotic Modes. *Electronics and Information Technologies*, 30. 137-154. https://doi.org/10.30970/eli.30.11

ABSTRACT

Background. Dynamic mode maps are a visualization tool used to analyze and classify the behavior of complex nonlinear dynamic systems as parameters change. They allow us to identify how the system's operating mode (steady state, periodicity, chaos, etc.) changes when one or more parameters are varied.

Materials and Methods. The paper proposes algorithms for constructing dynamic mode maps based on the convergence of periodicities and the minimum value function. The first is based on selecting the last element of the set and comparing this element in turn with all the previous ones. If the last element coincides with the previous one, then it is stated that the resulting set has a period of 1, which means that with these parameters, the system has a limit point. The second algorithm is based on creating arrays of standard deviations. Software has been developed for constructing dynamic mode maps using the convergence of periodicities and the minimum value function.

Results and Discussion. Based on the analysis of dynamic mode maps obtained by these methods under the condition of the existence of the Lifshitz invariant at $n = 3$, it was established that the method of periodicity convergence more fully describes the existing dynamics of the incommensurate superstructure, which is experimentally traced in tetramethylammonium tetrachlorocuprate crystals. It is shown that the dynamic mode maps calculated by the method of periodicity convergence have a significant number of existing periodicities and most fully describe the dynamics of the incommensurate superstructure. It is established that the palette of existing periodicities is more represented under the condition when the increment of the phase function acts as the arguments of the recurrent relations. The dynamic mode map deserves special attention when the increment of both the amplitude and phase functions acts as the arguments of the recurrent relations.

Conclusion. It has been established that the considered algorithm for constructing dynamic regime maps is effective for analyzing the dynamics of an incommensurate superstructure, which is described by a system of differential equations, and the appearance of an incommensurate superstructure is due to the existence of the Lifshitz invariant.

***Keywords*:** dynamical regime maps, incommensurate superstructure, incommensurate superstructure regimes, anisotropic interaction.

INTRODUCTION

Dynamic mode maps, which are diagrams on a plane with two parameters on the coordinate axes, provide a fairly complete and visual representation of a dynamic system`s behavior, showing the boundaries of areas of different dynamic modes. [1]. Constructed maps of dynamic modes are correlated with other ways of presenting this information, such

as a Lyapunov indicators map or Arnold tongues [2]. Both of these methods most often duplicate the information obtained by using the dynamic mode maps.

Two-dimensional mappings, like one-dimensional mappings, are given by recurrence relations of the form:

$$\begin{aligned} x_{n+1} &= f(x_n, y_n), \\ y_{n+1} &= g(x_n, y_n). \end{aligned} \tag{1}$$

In two-dimensional mappings, it will be necessary to deal with points on the plane, that is, with several numbers that specify the coordinates of the points. Two-dimensional mappings come into consideration in various ways. Some of them are the result of generalization of one-dimensional mappings, others model some phenomenon characterized by discrete time. Two-dimensional mappings often arise as difference schemes during the numerical solution of systems of differential equations.

Let's construct dynamic mode maps for an incommensurate superstructure described by two second-order differential equations [3]:

$$R'' - R^3 + \left(1 - \varphi'^2 + T\varphi'\right)R - R^{n-1}K(\cos n\varphi + 1) = 0, \tag{2}$$

$$\varphi'' + 2\frac{R'}{R}\left(\varphi' - \frac{T}{2}\right) + R^{n-2}K\sin n\varphi = 0, \tag{3}$$

where R, φ are dimensionless variables characterizing the amplitude and phase functions, T and K are dimensionless parameters, n is an integer characterizing the symmetry of the potential. In this system of differential equations, the appearance of an incommensurate superstructure is described by the Lifshitz invariant, T and K the parameters describe the long-range and anisotropic interaction, respectively. Anisotropic interaction in this system of differential equations is represented by the Dzyaloshynsky invariant [4].

The calculation of the spatial behavior of the amplitude and phase functions and the construction of dynamic mode maps is going to be performed in the Python software environment [5].

When constructing dynamic mode maps, a function describing the behavior of the disproportionation wave amplitude R or/and its change R' acted as the function $f(x_n, y_n)$. A function describing the behavior of the disproportionation wave phase φ or/and its change φ' acted as the function $g(x_n, y_n)$. The two-dimensional mapping constructed in this way depends on the values of parameters $a = K$ and $b = T$. The choice of these parameters K and T as a and b is due to the dependence of the non-commensurate superstructure dynamics on the parameter: T is the parameter that determines a long-range interaction magnitude; K is the parameter that determines an anisotropic interaction magnitude.

In this study, we are going to consider dynamic mode maps for the incommensurate superstructure that occurs in crystals of the tetramethylammonium tetrachlorometallate(II) group ($[N(CH_3)_4]_2MeCl_4$, where Me = Zn; Cu; Co; Fe; Ni), and its appearance is due to the presence of the Lifshitz invariant. The spatial variation of the order parameter amplitude and phase in these crystals can be described by expressions (2), (3). Dynamic mode maps will start to be observed, provided that the amplitude and phase functions, as well as their changes (first integral of the motion) will act as recurrence relations, in order to follow the full picture of disproportionate superstructure dynamics.

MATERIALS AND METHODS

Periodicity convergence method.

The following algorithm can be used to construct the dynamic mode maps.

Step 1. For some initial pair of values $(x_0; y_0)$ and some pair of values (a, b) according d_i to formula (1), a set of iterative values $(x_1; y_1), (x_2; y_2), \ldots, (x_n; y_n)$, is obtained.

Step 2. A few q last points are selected. For example, 8 or 16. q is a number that represents the maximum period that can be determined by this algorithm. Selected points form a set $(x_{n-q+1}; y_{n-q+1}), (x_{n-q+2}; y_{n-q+2}), \ldots, (x_{n-1}; y_{n-1}), (x_n; y_n)$.

Step 3. Working with a set of points that have 2 coordinates is not very convenient in our case. To do this, you need to replace 2 coordinates with some single number. For example, replace with the square of the modulus of the vector corresponding to the given coordinate. That is, instead of the point $(x_i; y_i)$, the value $d_i = x_i^2 + y_i^2$ is used. As a result, the set $d = (d_1, d_2, \ldots, d_q)$ is obtained. Further actions will require d_i displaying some specific $(x_i; y_i)$ and not allow repeats. But this cannot be achieved if the points form a certain symmetrical figure, for example, a circle – distances to the middle of the coordinates center of which will be the same. Therefore, it makes sense to introduce some asymmetry in the above operation. For example: $d_i = (x_i + k_1)^2 + (y_i + k_2)^2$.

Step 4. Work with a set d is the next. The last element of the set is highlighted and this element is compared one by one with all the previous ones. If the last q element matches $q - 1$, then it is claimed that the obtained set d has a period of 1, which means that the system has a limit point with the parameters (a, b). If the match is on $q - 2$, then it is believed that a cycle with a period of 2 is implemented in the system. And so on, until the period is not detected. That is, upon reaching a situation where none of the numbers in set d will match the last element. In this case, the presence of chaos is asserted. Although in reality it can turn out to be a period of higher order, or the system has not yet moved to stable behavior, and we ourselves have the right to choose what accuracy the behavior of the system is considered with. In a situation where the period was not detected, it is reasonable to record it as zero. The method of determining the period listed above can be called a *simple comparison* method or a *periodicity convergence* method.

Step 5. Next, it is necessary to save the obtained result. Values are stored in 3 pre-created arrays according to the following rule: in the first array – the value of a, in the second - the value of b, and finally, in the third – the value of the received period. Of course, serial numbers of the data must match.

Step 6. All the previous points are performed with other a and b values. It is necessary to walk through all possible combinations (a, b) from the interval with a certain step.

Step 7. The last step is displaying the image. Points corresponding to different period values must be displayed using different colors. Validation loops or logical operations can be used for this.

Method of minimum value function

However, there is also a slightly different approach to determining the periods that appear in the image. This algorithm repeats the above one except **Step 4.** The fact is that in the new algorithm, a special function that analyzes the array according to its algorithm will be used to determine the periodicity in the array d. Let's call this function ***Min_value*** and see how it works.

Step 1. The function ***Min_value*** accepts an array $d = (d_1, d_2, \ldots, d_n)$. The choice of number n is quite arbitrary, but for working with the function ***Min_value***, it will be very beneficial to take n = 8, 16 or 32. Taking a greater number does not make sense.

Step 2. Arrays of the following form are constructed from the array d: $d_1 = (d_1, d_1, \ldots, d_1)$, $d_2 = (d_1, d_2, d_1, d_2, \ldots, d_1, d_2)$, $d_3 = (d_1, d_2, d_3, \ldots)$, and so on, up to $d_{n/2} = (d_1, d_2, \ldots, d_{n/2}, d_1, d_2, \ldots, d_{n/2}, \ldots)$. That is, arrays consisting of the first

element d, the first two, the first three, and so on are formed. The length of all arrays is equal n.

Step 3. Next, an element-by-element difference between d and d_i is taken. For example, the difference between d and d_1: $\Delta d_1 = d - d_1 = (d_1 - d_1, d_2 - d_1, \ldots, d_n - d_1)$. After that, the standard deviation of the set Δd_1 is taken and written as λ_1 a number. And this is repeated with all Δd.

Step 4. An array is formed from the set $\lambda_1, \lambda_2, \ldots, \lambda_{n/2}$ and is presented as a result of a ***Min_value*** *function*.

The function ***Min_value*** outputs an array λ, and the closer λ_i is to zero, the more likely d has period i.

It is clear that the task is to select the smallest element from the set λ. However, it is more suitable to choose the first element i, which satisfies the condition $\lambda_i < \varepsilon$, where ε is a pre-set small value.

There are several remarks about the algorithms. The fact is that the systems used are non-linear, so sometimes values of points coordinates go beyond some reasonable limits, and the program writes the coordinates as (±inf, ±inf) or (NaN, NaN). However, it is better to exclude both cases from consideration and not to display them on the plane, because when comparing two inf values, the program will show that they are equal, and in our case, it is a period equal to one.

The second remark concerns accuracy. If a sufficient number of iterations is taken, there is no doubt that the system has arrived at its stable behavior mode, be it a periodic mode, a chaotic mode, or a stable point. In this case, it is not necessary to require the last element of the sequence to match with some previous elements exactly. It is enough that their difference is less than some given small ε value. In this case, it is possible to avoid various difficulties with "long convergent" sequences, which often arise when the period is equal to one.

The construction of dynamic mode maps is a task in itself and requires a fairly large amount of machine time, so when building the program, displaying the image inside the cycle in the period calculation should be avoided. In addition, during the initial debugging of the program, a small step of changing parameter values should not be used - the execution time of the program strongly depends on this [2].

The process of developing software code for calculating the dynamic mode map using the minimum value function is presented in the **Appendix**.

RESULTS AND DISCUSSION

Dynamic mode maps of incommensurate superstructures described by the Lifshitz invariant and obtained using the function of minimum values

Let's consider the constructed dynamic mode maps for crystals of group A_2BX_4 constructed according to the above methods.

Crystals of A_2BX_4 group, particularly tetramethylammonium tetrachlorometalates, are characterized by a complex sequence of phase transitions, including a transition to an incommensurate phase with nanoperiodicity (with a period ~ 100..160 nm). Such a wealth of commensurate (long-periodic) phases set and incommensurate phases sequence in these objects stimulated the study of dynamic mode maps of these systems.

These objects are characterized by different multiplications of elementary cell n. This parameter also characterizes the symmetry of the thermodynamic potential. That is, depending on the symmetry (multiplication of elementary cell n), one or another sequence of phase transitions will be observed in the crystal. The first representative of this family is the crystal of tetramethylammonium tetrachlorocuprate (TMATCC, chemical formula $[N(CH_3)_4]_2CuCl_4$), which in the ferroelastic phase is characterized by elementary

multiplication equal to n = 3. At the phase transition T_i = 297K in a special point of Brillouin zone (m/n, where n = 3, m = 1) an incommensurate superstructure with a wave vector $q = (1/3 - \delta)/c$ appears. Changes of δ value are insignificant ($\delta \approx 0.007..0.0075$). At the same time, the modulation wavelength near T_i is $L_i \approx 1617$ Å. With a decrease in temperature, down to the transition to a commensurate ferroelastic phase (T_c = 291 K), the process is accompanied by an increase in anisotropic interaction, which is described by the Dzyaloshynskyi invariant and determined by the parameter K. The value of T parameter describes the long-range interaction process and may decrease as the temperature in the incommensurate phase decreases. Therefore, the T and K parameters describe the dynamics of the disproportionate superstructure.

Fig. 1 and Fig. 2 show dynamic mode maps for the following parameters of the system under study: T = `c0` = 0.01..1.0, with a step of 0.01, `c2` = 3, T = 200, K = `c1` = 0.01..1.0, with a step of 0.01, provided n = 3

- `d.append((data2[i-1]+k1)**2+(data3[i-1]+k2)**2)` for Fig.1
- `d.append((data0[i-1]+k1)**2+(data1[i-1]+k1)**2+(data2[i-1]+k1)**2+(data3[i-1]+k2)**2)` for Fig.2

Visualizations made it possible to trace the dependence of existing periodicities on the system parameters values. If `c0` = 0.01..1.0, with a step of 0.01, `c2` = 3, T = 200, `c1` = 0.01..1.0, with a step of 0.01, then the existence of a cascade of periodicities is

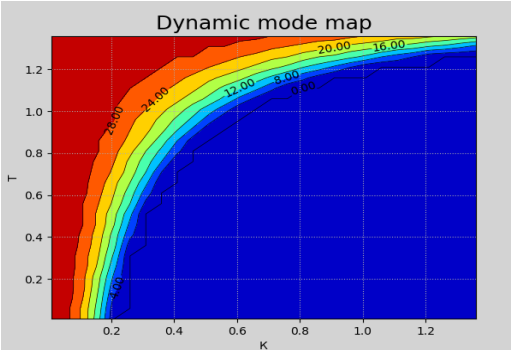


Fig. 1. The two-dimensional dynamic mode map is constructed using the function of minimum values for a system with the existence of a non-commensurate superstructure with n = 3, provided c0 = 0.01..1.4, with a step of 0.01, c2 = 3, c1 = 0.01..1.4, with a step of 0.01.

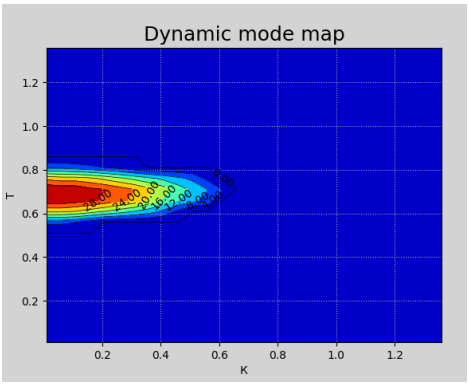


Fig. 2. The four-dimensional dynamic mode map is constructed using the function of minimum values for a system with the existence of a non-commensurate superstructure with n = 3, under the condition c0 = 0.01..1.4, with a step of 0.01, c2 = 3, c1 = 0.01..1.4, with a step of 0.01.

observed. This indicates the presence of stable and unstable system movement trajectories in different areas.

From conducted studies of the influence of system parameters `c0`, `c1` on existing modes, it was established that `c0` parameter determines the scale and number of existing periodicities, while `c1` is their interval. This allows us to control the surface shape by specifying different combinations of parameters.

According to Fig.1, areas with different values, where red zones indicate high values, which may indicate the existence of periodicities with a longer period, and blue zones - with a smaller value, which indicates more stable areas. This distribution demonstrates the complexity of system dynamics and possible transitions between different behavior modes.

The resulting dynamic mode map of the system under research described by the system of differential equations (1.2, 1.3) does not correspond to the experimentally observed modes obtained in [7] provided $n = 3$. Namely, the existence of two incommensurate structures and corresponding dependence of incommensurate superstructure modes on the amount of anisotropic interaction cannot be traced [8]. Therefore, let's consider the four-dimensional dynamic mode map, which is shown in Fig. 2. In Fig. 2, it is possible to observe a cluster of values in the upper part, closer to the central axis, which indicates an important role in appearance of periodicities in the magnitude of long-range interaction. Analysis of dynamic modes on such maps helps to understand where the system may be unstable and where it becomes stable, which is important for development of effective management and forecasting strategies.

The resulting four-dimensional dynamic mode map also does not fully describe the behavior of Lyapunov coefficients for systems with $n = 3$ [7,8]. Namely, if $T = 1$, then this system is characterized by the existence of two chaotic regimes, which are spatially separated by the magnitude of anisotropic interaction. The dynamic mode map shown in Fig. 2 is characterized by the existence of the cascade of periodicities in the vicinity of $T = 0.7$ and $K = 0..0.6$. An absence of the chaotic regime proves that this method of constructing dynamic mode maps is not perfect. Therefore, let's consider the construction of dynamic mode maps obtained using periodicity convergence method.

Dynamic mode maps of non-commensurate superstructures described by the Lifshitz invariant obtained using the periodicity convergence method, provided *n* = 3.

Fig. 3 shows dynamic mode maps depending on values of parameters $a = K$ and $b = T$ with $n = 3$ obtained using the periodicity convergence method. Dynamic mode maps are given, where amplitude (R) and phase (φ) functions of the non-interchangeable superstructure, and their changes (R' and φ' respectively) are the arguments of recurrence relations. Thus, Fig. 3 shows dynamic mode maps under the condition of consideration as function arguments: R, φ (Fig.3a); R', φ' (Fig.3b); R, φ' (Fig.3c); R', φ (Fig.3d).

Analyzing the obtained dynamic mode maps, it should be noted that a palette of existing periodicities is more represented if the increment of phase function (φ') acts as the argument of the recurrence relations. The dynamic mode map deserves special attention when increase in both amplitude (R') and phase (φ') functions acts as arguments of recurrence relations (Fig. 3b). This is not surprising, because for this system the Lyapunov coefficients, characterized by the largest changes, are described by spatial changes in the amplitude (R') and phase (φ') of the order parameter [4]. The given dynamic mode map in Fig. 3, b,c,e is characterized by a richer palette of colors corresponding to different periodicities. This is especially evident when the phase and its change act as arguments of recurrence relations (Fig. 3, e).

When constructing dynamic mode maps, we chose only the first eight possible periodicities (red - 1, orange - 2, yellow - 3, green - 4, cyan - 5, blue - 6, violet - 7, black - all others). The absence of any periodicities is marked in white. Periodicities with the

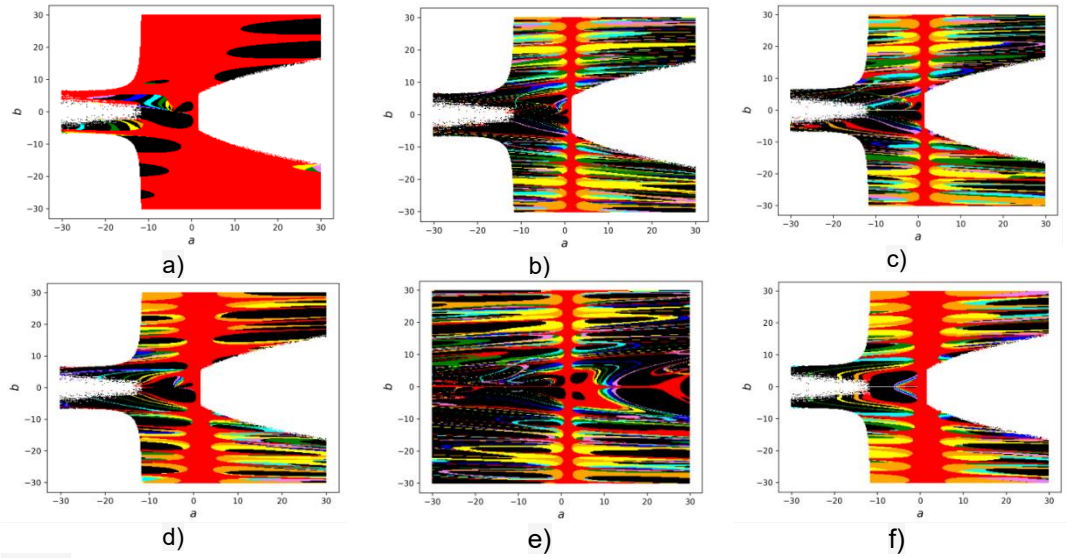


Fig. 3. Dynamic regime map along axes a and b. Here, $a = K$ represents the value of the anisotropic interaction parameter described by the Dzyaloshinskii invariant, and $b = T$ corresponds to the value of the long-range interaction parameter of the system, under the condition $n = 3$, with variations of the following pairs of variables: a) R, φ; b) R', φ'; c) R, φ'; d) R', φ; e) φ, φ'; f) R, R'.

smallest period (red - 1; orange - 2; yellow - 3; green - 4) are dominant on the dynamic mode maps.

Based on the dependence of the value of Lyapunov coefficients [4] on anisotropic interaction value (K), if $n = 3$, and $K > 1.0$, then a transition to a chaotic state is observed. Therefore, Fig. 4 shows dynamic mode maps when the parameters $a = K$ and $b = T$ are changed in the range from –2 to 2.

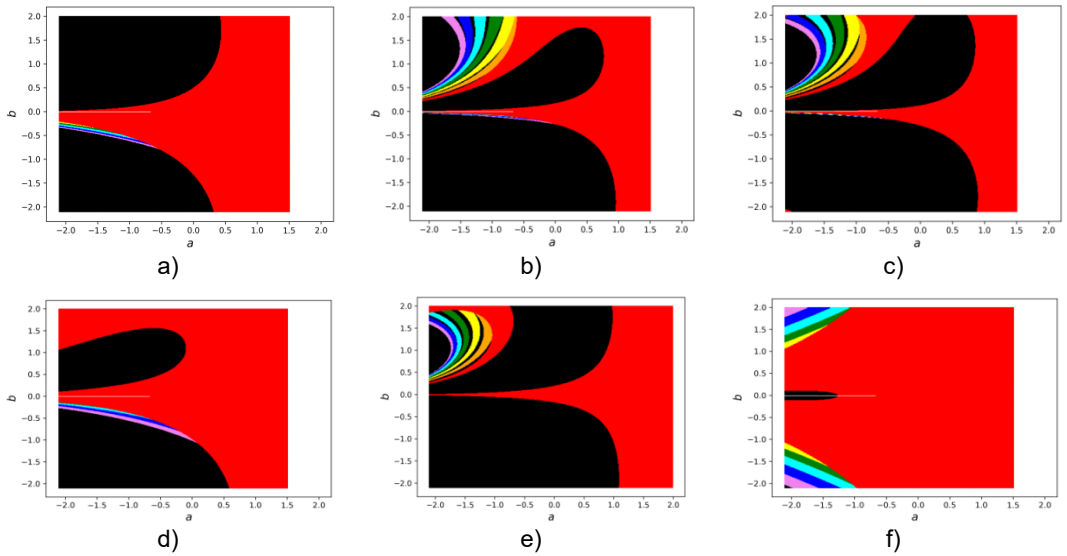


Fig. 4. Dynamic regime map along axes a and b. Here, $a = K$ denotes the value of the anisotropic interaction parameter described by the Dzyaloshinskii invariant, and $b = T$ represents the value of the long-range interaction parameter of the system, under the condition $n = 3$, $\varepsilon = 0.1$, and variations of the following pairs of variables: a) R, φ; b) R', φ'; c) R, φ'; d) R', φ; e) φ, φ'; f) R, R'.

The dynamic mode maps shown in Fig. 4 are similar to those shown in Fig. 3. For these dynamic mode maps, the same regularity is observed. Namely, dynamic mode maps are characterized by a greater number of periodicities, provided that the phase increase value (φ') acts as the argument of recurrence relations. When the amplitude value or its increase value acts as an argument of recurrence relations, then the dynamic mode maps show a sharp transition to a state characterized by an increase in the amplitude function value and its change ($K > 1.5$). It is known [8] that such an increase in the amplitude of spontaneous deformation in TMATCC crystals is associated with a transition to a commensurate ferroelastic phase. At the same time, spatial incommensurate modulation disappears. It should also be noted that the incommensurate phase in the TMATCC crystals is one, but its structure is different (I_1 and I_2) in different areas of the phase diagram.

Since on dynamic mode maps black color corresponds to periodicities with a period greater than $N = 8$, then periodicities with periods greater than this value ($N > 8$) may exist in this spatial area. Therefore, this area can be considered as an area of possible coexistence of incommensurate phases with different structures. Exiting the black area and transitioning to the red area indicates the transition to the first periodicity. It can be assumed that this transition is associated with a change in the superstructure mode from sinusoidal to soliton. But, as is known, the soliton mode is characterized by wave harmonics of incommensurate modulation caused by the anharmonicity of oscillations of structural units, as well as the existence of anisotropic interaction. That is, the spectrum of existing oscillations is richer than in the sinusoidal mode. Therefore, the reason for such an obtained result (Fig. 4.) may be in the choice of ε value. The value of the ε variable indicates the permissible difference between d_1 values and consecutive values $d_2 ... d_8$. When calculating the mapping, only the last 8 values of d are determined, which depend on x, y coordinates:

$$d_i = (x_i + m_1) + (y_i + m_2),$$

where m_1 and m_2 are constants necessary to introduce some asymmetry.

If the difference $|d_1 - d_2| < \varepsilon$, then we consider that the cycle is equal to one iteration and the corresponding dot is colored in the first color (red). If the cycle is equal to two iterations, then the color of the dot is orange, etc. For the cycle equal to 7, the dot color is purple. If the cycle is greater than 7, then we consider that the transition between iterations is chaotic, and such a point is colored black. If d goes to infinity, the calculation of the iteration cycle is interrupted, and the dot color is considered white. That is, the ε variable essentially determines the existence of periodicities with the value ε. So, returning to the appearance of periodicity, which is determined by the red color when $K > 1$ (Fig. 4, b,c,e), according to the authors, it can be associated with the non-optimized value of the ε variable. For this purpose, a study of the effect of ε value on the dynamic mode maps character was carried out. Since the studied system, which determines the appearance and dynamics of the incommensurate superstructure, is described by a system of second-order differential equations, let's consider a four-dimensional mapping. The amplitude function (R), its spatial variation (R'), the phase function (φ), and its spatial variation (φ') are the arguments of the recurrence relations. Fig. 5 shows four-dimensional dynamic mode maps for different ε values, provided that only the last 14 values of d are determined (that is, an increase in the number of possible periodicities from 8 to 16). Based on the analysis of obtained dynamic mode maps (Fig. 5), it follows that at $\varepsilon = 0.01$, under the condition that (T) parameter describing the long-range interaction is equal to $a = T = 1$, the consistency between the behavior of Lyapunov parameters is observed under the condition of changing the anisotropic interaction $b = K$. Namely, the existence of two "black" areas describing the presence of two incommensurate phases, and disappearance of the incommensurate phase at $K = 1.5$. At other values of ε ($\varepsilon = 1$ Fig. 5d; $\varepsilon = 0.1$ Fig. 5a; $\varepsilon = 0.001$ Fig. 5c),

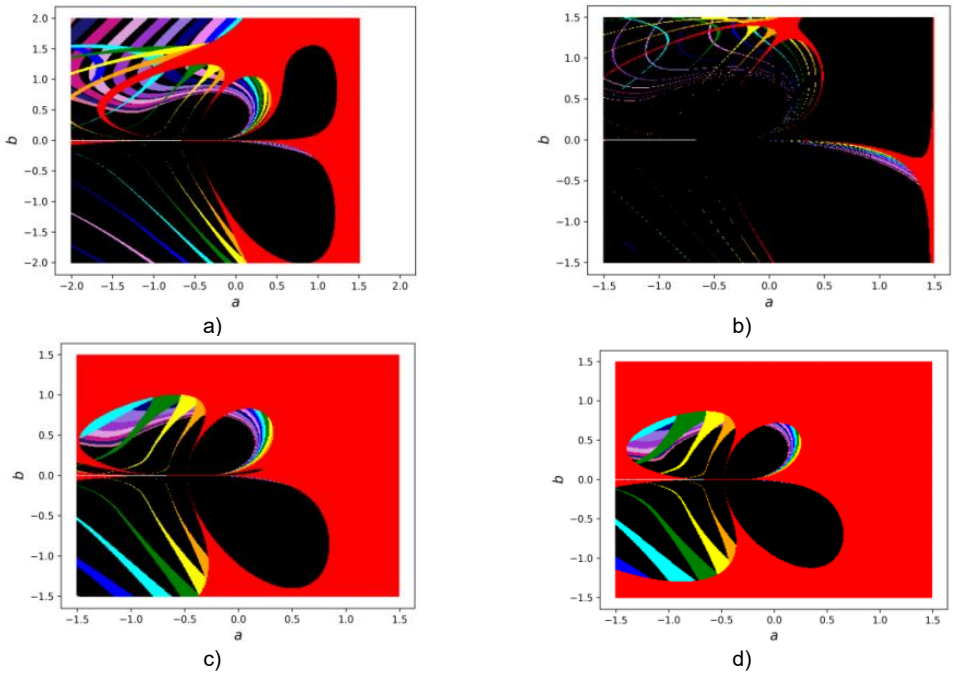


Fig. 5. Four-dimensional dynamic regime maps along axes a and b. Here, $a = K$ represents the value of the anisotropic interaction parameter described by the Dzyaloshinskii invariant, and $b = T$ denotes the value of the long-range interaction parameter of the system, under the condition $n = 3$, with varying values of ε: a) $\varepsilon = 0.1$; b) $\varepsilon = 0.01$; c) $\varepsilon = 0.001$; d) $\varepsilon = 1.0$.

dynamic mode maps are characterized by the existence of narrowed spatial areas that describe chaotic states, that is, existing incommensurate phases, and by expanding the spatial areas of periodicities with lower periods. Since the existing periodicities can be associated with localizations of the wave incommensurability vector on commensurate high-order values, then the presence of narrow intervals of their existence (periodicities) on the dynamic mode maps is in favor of the variable $\varepsilon = 0.01$.

Fig. 6 shows a map of the existence of possible periodicities from the anisotropic interaction value $a = K$, under the condition of the maximum long-range interaction value $T = 1$. According to Fig. 6, an increase in the ε variable leads to a decrease in the intervals of existence of periodicities. It should also be noted that two intervals (by K) of the existence of chaotic states (incommensurate phases), which are characterized by different periodicity

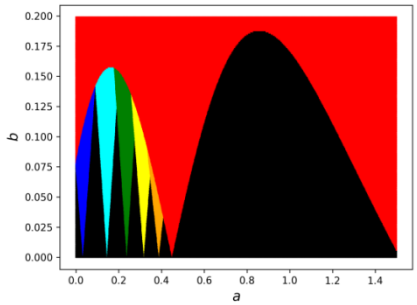


Fig. 6. Two-dimensional map of the existence of possible periodicities. Where $a = K$ is the anisotropic interaction parameter value, which is described by the Dzyaloshynsky invariant, $b = \varepsilon$, provided that $n = 3$ and the maximum value of the long-range interaction $T = 1$.

dynamics, are observed. The first interval ($K = 0..0.4$) is characterized by the existence of periodicities both with $N < 14$ and with $N > 14$, and the other interval ($K = 0.4..1.5$) is characterized by periodicities with $N > 14$.

Proceeding from such dynamics of the existing periodicities from the anisotropic interaction parameter, it can be assumed that there is one incommensurate phase in this crystal, and at $K = 0.4$ there is only a change of incommensurate superstructure mode from sinusoidal to soliton. Based on Fourier studies of spectra of amplitude and phase functions of the incommensurate superstructure [7], the soliton mode is described by the existence of several periodicities. Dependence of harmonics of a wave vector (q) of incommensurate modulation on the amount of anisotropic interaction (parameter K), described by the Dzyaloshynskyi invariant, indicates complex transformations that occur during transition from sinusoidal mode to soliton mode at $K = 0.4$. At the same time, both the disappearance of existing periodicities and the appearance of new ones were traced.

To find out processes that occur at $K = 0.4$, let's consider two-dimensional dynamic mode maps provided that $\varepsilon = 0.01$ (Fig. 7). According to Fig. 7, the transition from one chaotic state to another state passes through a cascade of periodicities. These periodicities determine the peculiarity of the first disproportionate area behavior (a chaotic state in the vicinity $K = 0..0.5$).

It is known that as a result of the appearance of the incommensurate superstructure wave oscillations harmonic, it passes from the sinusoidal mode to the soliton mode. If it is assumed that in the sinusoidal mode, harmonics of the incommensurate superstructure, which causes its transition to the soliton mode, arise, then the range of sinusoidal mode and the mode transitional to the soliton mode should be characterized by the existence of periodicities, and in the ideal case, the soliton mode should have at least as many periodicities. The observed "black" area with increasing K value ($K = 0.5..1.5$), in the absence of any periodicities with $N < 14$, is possibly related to the formation of a chaotic superstructure. This is particularly indicated by the dependence of Lyapunov coefficients [3] on the K parameter. The appearance of the incommensurate superstructure chaotic

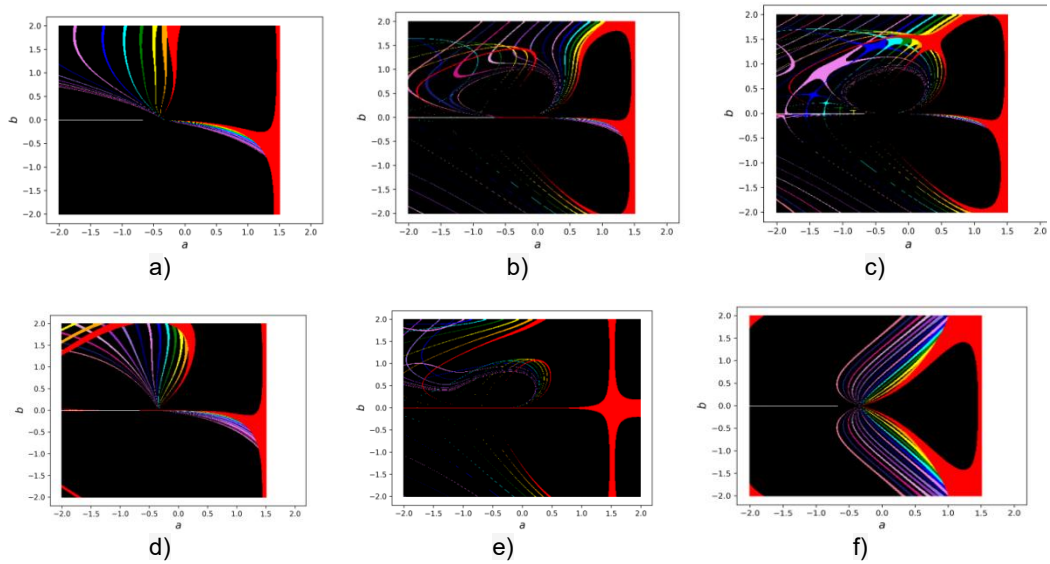


Fig. 7. Dynamic regime map along the axes *a* and *b*. Here, $a = K$ represents the value of the anisotropic interaction parameter described by the Dzyaloshinskii invariant, and $b = T$ denotes the value of the long-range interaction parameter of the system, under the conditions $n = 3$, $\varepsilon = 0.01$, and variations in the following pairs of variables: a) R, φ; b) R', φ'; c) R, φ'; d) R', φ; e) φ, φ'; f) R, R'.

state on these dependencies begins to manifest itself at $K \geq 1.0$. Taking into account the dynamic mode maps in Fig. 7, the transition to a chaotic state is caused by a change in the phase function of the incommensurate superstructure. The amplitude function is responsible for the presence of periodicities in the first chaotic phase.

CONCLUSION

Therefore, the considered algorithm for constructing dynamic mode maps is effective for analyzing the incommensurate superstructure dynamics, which is described by a system of differential equations, and the appearance of an incommensurate superstructure is due to Lifshitz invariant. The software for constructing dynamic mode maps using the periodicity convergence method and the function of minimum values was developed. Based on the analysis of dynamic mode maps obtained by these methods under the condition of Lifshitz invariant existence at $n = 3$, it was established that the periodicity convergence method more fully describes the existing incommensurate superstructure dynamics, which is experimentally observed in TMATCC crystals. It was shown that the dynamic mode maps calculated by the periodicity convergence method have a significant number of existing periodicities and most fully describe the dynamics of the incommensurate superstructure. It was confirmed that the incommensurate phase in the TMATCC crystals is single, but its structure is different (I_1 and I_2) in various phase diagram areas.

COMPLIANCE WITH ETHICAL STANDARDS

The authors declare that they have no competing interests.

AUTHOR CONTRIBUTIONS

Conceptualization, [S.S.]; formal analysis, [S.S., I.Ka., Y.S.,].; investigation, [I.Ku.]; resources, [I.Ku.]; writing – original draft preparation, [S.S., I.Ka., I.Ku.]; writing – review and editing, [I.Ka, I.Ku, Y.S.].

All authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] Strogatz, S. H. (2024). *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. Chapman and Hall/CRC.

[2] Pichkur V.V., Kapustyan O.V., Sobchuk V.V. (2020) *Theory of dynamic systems*. Lutsk: Vezha-Druk, 348 p. (in Ukrainian) URL: https://mechmat.knu.ua/wp-content/uploads/2023/01/teoriia_dynamichnykh_system_kapustianov_pichkurvv_sobchukvv.pdf

[3] Cummins, H. Z. (1990). Experimental studies of structurally incommensurate crystal phases. *Physics Reports*, *185*(5-6), 211-409. https://www.sciencedirect.com/science/article/abs/pii/037015739090058A

[4] Bak, P. (1981). Chaotic behavior and incommensurate phases in the anisotropic Ising model with competing interactions. *Physical Review Letters*, *46*(13), 791. https://doi.org/10.1103/PhysRevLett.46.791

[5] Andersson, C., Åkesson, J., & Führer, C. (2016). Pyfmi: A python package for simulation of coupled dynamic models with the functional mock-up interface. *Technical Report in Mathematical Sciences*, (2). ISSN 1403-9338. PyFMI.pdf

[6] Lynch, S. (2018). *Dynamical systems with applications using python*. Basel, Switzerland: Birkhäuser. 665 p. https://doi.org/10.1007/978-3-319-78145-7

[7] Kunyo, I. M., Karpa, I. V., Sveleba, S. A., & Katerinchuk, I. M. (2019). Dimensional effects in dielectric crystals [N(CH3)4]2MeCl4 (Me= Cu, Zn, Mn, Co) with

incommensurate phase: monograph. *Lviv: Ivan Franko Lviv National University*, *220*. (in Ukrainian).
[8] Sveleba, S., Katerynchuk, I., Kuno, I., Shmygelsky, Y., Velgosh, S., & Bukhnii, Y. (2024) Dynamics of incommensurable superstructure in crystals with thermodynamic potential symmetry n=4. *Electronics and information technologies/Електроніка та інформаційні технології*, (27), P.132. (in Ukrainian). https://doi.org/10.30970/eli.27.10.

APPENDIX

Development of a program for calculating the dynamic mode map using the minimum values function.

This source code is intended for modeling dynamic systems and studying the dependence of the order parameter amplitude and phase on the system parameters. It uses numerical methods to solve systems of differential equations and visualize the results. The main parameters analyzed include long-range interaction (parameter c_0) and anisotropic interaction (parameter c_1). Influence of these parameters is studied through an integration of the system of equations with time. [6]

1. **Import libraries**:

Libraries provide functionality for calculating differential equations, data visualization, and numerical calculations.

```
from jitcdde import y, t
from symengine import sin, cos
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
import numpy as np
from scipy.integrate import ode
from scipy.integrate import odeint
import array as arr
```

1. **Definition of initial parameters**:

- c_0 = 1.0 – long-range interaction
- c_1 = 0.1 – anisotropic interaction
- c_2 = 3 – symmetry of thermodynamic potential
- T = 1000, dt = 0.1

Initial values for system parameters.

```
y0 = [0.3, 0.0, 0.0, 0.0]
```

2. **Definition of differential function:**

```
def f(t, y):
 y0, y1, y2, y3 = y
 return [y1, y0**3 - (1 + c0 * y3 - y3**2) * y0 + c1 * (y0**(c2 - 1)) * (1 +
np.cos(c2 * y2)), y3, -(y1 / y0) * (2 * y3 - c0) - c1 * (y0**(c2 - 2)) *
np.sin(c2 * y2)]
```

This function defines a system of differential equations which describe the dynamics of the system under study (given by expressions (1.2) and (1.3)). Nonlinear terms, long-range interaction c_0 and anisotropic interaction c_1 are taken into account here.

3. **Initialization of arrays for saving data:**

Arrays are used to store data at various stages of calculations.

```
cz=arr.array('d',[])
y=arr.array('d',[])
```

```
t=arr.array('d',[])
d=arr.array('d',[])
aT=arr.array('d',[])
bK=arr.array('d',[])
H=arr.array('d',[])
```

4. Numerical integration of differential equations:

```
params = np.arange(0.01, 1.41, 0.01)
for c0 in np.arange(0.01, 1.01, 0.01):
 for c1 in params:
 r = ode(f).set_integrator("vode", method="adams", order=10, rtol=0, atol=1e-
8, with_jacobian=False)
 y0 = [0.3, 0.0, 0.0, 0.0]
 r.set_initial_value(y0, 0)
 T = 200
 dt = 0.1
 t = []
 y = []
 while r.successful() and r.t <= T:
 r.integrate(r.t + dt)
 y.append(r.y)
 t.append(r.t)
 t = np.array(t)
 y = np.array(y)
```

Numerical integration of differential equations is performed using the Adams method.

5. Analysis of results:

After integration, the data is analyzed to determine the frequency and amplitude characteristics. The change of phase and amplitude of the order parameter is calculated for each combination of parameters.

```
n=len(t)
y.shape=n,4
data1 = y[ :,1]
data0 = y[ :,0]
data3 = y[ :,3]
data2 = y[ :,2]
nn=len(data0)
x=data0
xx=data1
y=data2
yy=data3
d=[]
```

6. Calculation of standard deviation:

```
def Min_value (d):
 n = 32
 g = np.zeros((n, n))
 for i in range(1, int(n / 2)):
 p = d[1:i]
 for k in range(1, int(np.fix(n / i) + 1)):
 try:
 if (i + k * i) <= n:
 g[(k-1)*i+1:k*i, i] = d[k*i+1:i+k*i] = p
 else:
 g[(k-1)*i+1:(k-1)*i+n-k*i, i] = d[k*i+1:n] - p[1:n-k*i]
 except ValueError:
 continue
```

```
    return round(np.std(g), 7)
```

Min_value function calculates the standard deviation for the data used to analyze the stability of the dynamic system.

7. Data processing:

```
aT=np.array(aT)
bK=np.array(bK)
H=np.array(H)
eps=0.01
p3=[]
LL=sorted(set(H))
for i in range(len(LL)):
 for j in range(len(H)):
 if H[j]==LL[i]:
 if i<=16:
 H[j]=16-i
 if i>16:
 H[j]=0
p=np.arange(0.01,1.41,0.01)
q=np.arange(0.01,1.01,0.01)
l=len(p)
j=len(q)
x=aT
y=bK
print("j size", j)
print("l size", l)
H.shape=j,l
```

This code snippet converts **aT**, **bK** and **H** arrays to **numpy.array** type for further processing. The **eps** value is set to 0.01. The empty list **p3** is initialized. Then unique values from the **H** array are sorted and result is stored in the **LL** variable. For each value in **LL,** a double iteration on all elements of the **H** array is performed, comparing them with the current value from **LL**. If the values match, they are modified: for the first 16 values of index **i,** the value of **H** is reduced by **16 - i**, for other values, **H** is set to 0.

Arrays of **p** and **q** parameters are created, where **p** contains values from 0.01 to 1.41 with a step of 0.01, and **q** contains values from 0.01 to 1.01 with a step of 0.01. Lengths of these arrays are determined and stored in variables **i** and **j**. Variables **x** and **y** are assigned the values of the arrays **aT** and **bK** respectively. Sizes of **j** and **i** arrays are displayed on the screen for verification. Finally, the **H** array changes its shape according to **j** and **i** values, which is necessary for the correct data displaying in the form of 2D or 3D graphs.

8. Visualization of results:

```
fig = plt.figure()
ax = Axes3D(fig)
u = np.array(q)
Y, X = np.meshgrid(p, q)
z0 = H
contour = plt.contour(X, Y, z0)
plt.clabel(contour, colors = 'k', fmt = '%2.1f', fontsize=20)
c = ('#ff0000', '#ffff00', '#0000FF', '0.6', 'c', 'm')
contour_filled = plt.contourf(X, Y, z0, colors=c)
plt.colorbar(contour_filled)
plt.title('Dynamic mode map')
plt.xlabel(u'K')
plt.ylabel(u'T')
plt.savefig('R`,f`_32,eps=0.001,n=4_A.png', dpi=300)
plt.show()
```

Visualization of the results includes construction of contour diagrams and 3D graphs showing the dependence of dynamic modes on *K* (anisotropic interaction) and *T* (long-range interaction) parameters. This makes it possible to evaluate the influence of the parameters on system behavior and to determine areas of stability and instability.

In this code, parameters have the following values and interpretations:

- **T = 1000**: Total integration time - a time limit up to which the numerical integration of differential equations is performed.
- **dt = 0.1**: Time step for the numerical integration - an interval between points where values of the system variables are calculated.
- **n = 32**: Number of first selected periodicities.

1.2. Construction of dynamic mode maps using periodicity convergence method.

Let's look at the code that describes the method of conducting experiments to study the dependence of order parameter phase and amplitude on parameters of long-range interaction (c0) and anisotropic interaction (c1).

1. **Import libraries**:

The libraries provide functionality for working with differential equations, data visualization, and numerical calculations.

```
from jitcdde import y, t
from symengine import sin, cos
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
import numpy as np
from scipy.integrate import ode
from scipy.integrate import odeint
import array as arr
```

2. **Definition of initial parameters**:

- c0 = 1.0 – long-range interaction
- c1 = 0.1 – anisotropic interaction
- c2 = 3 – variable parameter
- T = 1000, dt = 0.1

Initial values for system parameters.

```
y0 = [0.3, 0.0, 0.0, 0.0]
```

3. **Definition of differential function**:

```
def f(t, y):
 y0, y1, y2, y3 = y
 return [y1, y0**3 - (1 + c0 * y3 - y3**2) * y0 + c1 * (y0**(c2 - 1)) * (1 +
np.cos(c2 * y2)), y3, -(y1 / y0) * (2 * y3 - c0) - c1 * (y0**(c2 - 2)) *
np.sin(c2 * y2)]
```

The function defines a system of differential equations that describe the system dynamics taking into account nonlinear terms, long-range interaction (c0) and anisotropic interaction (c1).

4. **Initialization of arrays to save data**:

```
from scipy.integrate import ode
data=0.3
y=arr.array('d',[])
x=arr.array('d',[])
amplid=arr.array('d',[])
```

```
omega=arr.array('d',[])
chastota=arr.array('d',[])
chast1=arr.array('d',[])
d=arr.array('d',[])
crok=0.01
e=0.01
px=arr.array('d',[])
py=arr.array('d',[])
pc=[]
```

Arrays are used to store data at various stages of calculations.

5. **Numerical integration of differential equations**:

```
def Dynamic(c0, c1, e):
 d = []
 r = ode(f).set_integrator("vode", method="adams", order=10, rtol=0, atol=1e-
8, with_jacobian=False)
 y0 = [0.3, 0.0, 0.0, 0.0]
 r.set_initial_value(y0, 10)
 T = 100
 dt = 0.1
 t = []
 y = []
 while r.successful() and r.t <= T:
 r.integrate(r.t + dt)
 y.append(r.y)
 t.append(r.t)
 t = np.array(t)
 y = np.array(y)
 data0 = y[:, 0]
 data1 = y[:, 1]
 data2 = y[:, 2]
 data3 = y[:, 3]
 nn = len(data0)
 x = data0
 xx = data1
 y = data2
 yy = data3
 for i in range(len(data0)):
 d.append((x[i]+1)**2 + (xx[i]+1)**2 + (y[i]+1)**2 + (yy[i]+1)**2)
 c = 'black'
 try:
 if abs(d[14] - d[13]) < e: c = 'red'
 elif abs(d[14] - d[12]) < e: c = 'orange'
 elif abs(d[14] - d[11]) < e: c = 'yellow'
 elif abs(d[14] - d[10]) < e: c = 'green'
 elif abs(d[14] - d[9]) < e: c = 'cyan'
 elif abs(d[14] - d[8]) < e: c = 'blue'
 elif abs(d[14] - d[7]) < e: c = 'violet'
 elif abs(d[14] - d[6]) < e: c = '#000080'
 elif abs(d[14] - d[5]) < e: c = '#9370DB'
 elif abs(d[14] - d[4]) < e: c = '#9932CC'
 elif abs(d[14] - d[3]) < e: c = '#DDA0DD'
 elif abs(d[14] - d[2]) < e: c = '#C71585'
 elif abs(d[14] - d[1]) < e: c = '#191970'
 elif abs(d[14] - d[0]) < e: c = '#DB7093'
 except IndexError:
 pass
```

```
    return c
```

Numerical integration of differential equations is performed by the Adams method, using parameters c0 and c1. The Dynamic(c0, c1, e) function analyzes the system behavior and determines colors for visualization based on the calculated data.

6. **Visualization of results**:

```
for e in np.arange(0.0, 1.0, 0.01):
 for c1 in np.arange(-1.0, 1.5, 0.01):
 c = Dynamic(c0, c1, e)
 px.append(e)
 py.append(c1)
 pc.append(c)
plt.scatter(py, px, marker=',', s=0.2, color=pc)
plt.xlabel('$a$', fontsize=14)
plt.ylabel('$b$', fontsize=14)
plt.show()
```

Visualization of the results includes the construction of graphs showing the dependence of order parameter amplitude and phase on the magnitude of long-range interaction (c0) and anisotropic interaction (c1). Color coding is used to display different modes of dynamic behavior of the system.

Parameters and their interpretations

- **T = 1000**: Total integration time - a time limit up to which the numerical integration of differential equations is performed.
- **dt = 0.1**: Time step for the numerical integration - an interval between the points where values of the system variables are calculated.

This code allows the analysis of dynamic systems, taking into account long-range and anisotropic interactions. Using numerical integration and visualization, it is possible to construct dynamic mode maps and investigate the influence of parameters on system behavior.

ПРОГРАМИ РОЗРАХУНКУ КАРТИ ДИНАМІЧНИХ РЕЖИМІВ НА ПРИКЛАДІ СИСТЕМИ, ЩО ВОЛОДІЄ ХАОТИЧНИМИ РЕЖИМАМИ

Сергій Свелеба**, Іван Катеринчук, Іван Куньо, Ярослав Шмигельський
Львівський національний університет імені Івана Франка
вул. Ген. Тарнавського, 107, 79017 м. Львів, Україна

АНОТАЦІЯ

Вступ. Карти динамічних режимів – це інструмент візуалізації, який використовується для аналізу та класифікації поведінки складних нелінійних динамічних систем при зміні параметрів. Вони дозволяють нам визначити, як змінюється режим роботи системи (стаціонарний стан, періодичність, хаос тощо) при зміні одного або кількох параметрів.

Матеріали та методи. У статті запропоновано алгоритми побудови карт динамічних режимів на основі збіжності періодичностей та функції мінімального значення. Перший базується на виборі останнього елемента множини та порівнянні цього елемента по черзі з усіма попередніми. Якщо останній елемент збігається з попереднім, то стверджується, що результуюча множина має період 1, що означає, що

з цими параметрами система має граничну точку. Другий алгоритм базується на створенні масивів стандартних відхилень. Розроблено програмне забезпечення для побудови карт динамічних режимів з використанням збіжності періодичностей та функції мінімального значення.

Результати. На основі аналізу динамічних карт мод, отриманих цими методами за умови існування інваріанта Ліфшиця при $n = 3$, було встановлено, що метод збіжності періодичності повніше описує існуючу динаміку неспівмірної надструктури, яка експериментально простежується в кристалах тетрахлоркупрату тетраметиламонію. Показано, що динамічні карти мод, розраховані методом збіжності періодичності, мають значну кількість існуючих періодичностей та найповніше описують динаміку неспівмірної надструктури. Встановлено, що палітра існуючих періодичностей більш представлена за умови, коли приріст фазової функції виступає аргументами рекурентних співвідношень. На особливу увагу заслуговує карта динамічних мод, коли приріст як амплітудної, так і фазової функцій виступає аргументами рекурентних співвідношень.

Висновки. Встановлено, що розглянутий алгоритм побудови карт динамічного режиму є ефективним для аналізу динаміки неспівмірної надструктури, яка описується системою диференціальних рівнянь, а поява неспівмірної надструктури зумовлена існуванням інваріанта Ліфшиця.

Ключові слова: карти динамічних режимів, неспівмірна надструктура, режими неспівмірних надструктур, анізотропна взаємодія

Received / Одержано 08 May, 2025 | Revised / Доопрацьовано 05 June, 2025 | Accepted / Прийнято 09 June, 2025 | Published / Опубліковано 26 June, 2025

Electronics and information technologies, 2025, 30, 155–162
http://publications.lnu.edu.ua/collections/index.php/electronics/index

UDC 004.932

INFLUENCE OF PREPARATION CONDITIONS ON THE ELECTRICAL CONDUCTIVITY OF GaN THIN FILMS

Oleh Bordun**, *Ihor Kukharskyy*, *Mariia Protsak*, *Ivanna Medvid*, *Iryna Kofliuk*, *Zhanetta Tsapovska
Ivan Franko National University of Lviv,
50 Dragomanov Str., 79005 Lviv, Ukraine

Bordun O., Kukharskyy I., Protsak M., Medvid I., Kofliuk I., Tsapovska Zh. (2025). Influence of Preparation Conditions on the Electrical Conductivity of GaN Thin Films. *Electronics and Information Technologies*, 30, 155-162. https://doi.org/10.30970/eli.30.12

ABSTRACT

Background. Gallium nitride (GaN) is a promising material for the developing of LEDs, lasers, ultraviolet detectors, and high-frequency transistors. The properties of GaN significantly depend on the type of defects and the parameters of thin film production. In particular, substrates, buffer layers, sputtering temperature, and partial pressure of the working gas have an impact. This led to studying the electrical conductivity of GaN thin films obtained by radio-frequency (RF) ion-plasma sputtering under different technological conditions.

Materials and Methods. GaN films with a thickness of 0.3..1.0 μm were deposited on sapphire and quartz substrates, with/without buffer layers of $MgAl_2O_4$ or ZnO. RF sputtering was carried out in a N_2 atmosphere at a pressure of 5×10^{-3} .. 5×10^{-2} Torr and temperatures of 400-650 °C. The structure was analyzed by X-ray diffraction, and the electrical conductivity was determined in the temperature range of 100..450 K by the method of two-point contacts using ohmic carbon contacts.

Results and Discussion. It was found that GaN films have a polycrystalline structure with different orientations, depending on the type of substrate and the presence of a buffer layer. The resistivity varies from 10^5 to 10^{10} ohm-cm. High-impedance samples are characterized by a thermal activation energy of 0.34 eV, and the donor type of conductivity in the films has been established. In the films on quartz substrates with a ZnO sublayer, lower activation energy values (0.004..0.05 eV) were recorded, indicating a different conduction mechanism. The conductivity of GaN increases with increasing substrate temperature and decreasing partial pressure of N_2 in the sputtering atmosphere, which leads to the formation of nitrogen vacancies – the main donor defects.

Conclusion. The conditions of their preparation largely determine the electrical conductivity of GaN thin films. The most important factors are the type of substrate, buffer layers, sputtering temperature, and gas pressure in the sputtering atmosphere. The obtained dependences of the change in the electrical conductivity of GaN thin films on the partial pressure of the working nitrogen gas and the temperature of the substrate during sputtering show that the most likely defects of the donor type in gallium nitride are nitrogen vacancies V_N.

Keywords: thin films, gallium nitride, electrical conductivity, donor defects.

INTRODUCTION

Gallium nitride (GaN) and compounds based on it are a promising material for the production of semiconductor light diodes and lasers emitting in the blue and ultraviolet regions of the spectrum, as well as ultraviolet radiation detectors [1-3]. Powerful high-speed

transistors based on GaN, which are superior to their silicon counterparts in terms of operating frequencies and power densities, have also become widely used [4, 5]. At the same time, an important task is to obtain GaN thin films with the required composition of intrinsic defects. Intrinsic atomic defects largely determine the optical and electrical properties of gallium nitride, such as absorption and emission spectra, type and value of electrical conductivity. Another important task is to find optimal substrates that match GaN in terms of lattice parameter and thermal expansion coefficient.

The basic task in the development of technology for electronic or optoelectronic devices based on gallium nitride is to obtain homogeneous thin film layers of high quality. One of the ways to solve this problem is to grow epitaxial GaN layers on sapphire substrates by using thin buffer layers. In this case, it is necessary to match the characteristics of the deposited buffer layer with the modes of high-temperature growth of the main GaN layer. The properties of the buffer layer in such a system are determined only indirectly, through the properties of the thin film layer deposited on it. On the other hand, the optimization of the deposition regime of the corresponding films is possible only if a suitable buffer layer is available. On suboptimal buffer layers, GaN thin films have poor quality regardless of the deposition method. Therefore, the procedure for determining the optimal parameters for depositing gallium nitride thin films is complex and complex. For this purpose, the electrical conductivity of thin films on sapphire and quartz substrates with buffer sublayers of $MgAl_2O_4$ and ZnO was studied in this work. The films were obtained by radio-frequency (RF) ion-plasma sputtering, which is considered optimal to produce semiconductor and dielectric films [6].

MATERIALS AND METHODS

Thin films of gallium nitride with a thickness of 0.3..1.0 μm were obtained by RF ion-plasma sputtering on sapphire substrates (Al_2O_3) and quartz substrates (SiO_2). In some cases, buffer sublayers of $MgAl_2O_4$ and ZnO were used. The RF sputtering was carried out in a nitrogen atmosphere at pressures from 5×10^{-3} to 5×10^{-2} Torr. The target for sputtering was metallic Ga. The temperature of the substrates during sputtering varied from 400 to 650 °C, and the RF discharge power was from 100 to 150 W.

The structure and phase composition of the obtained films were studied by X-ray diffraction analysis (Shimadzu XDR-600). The characteristic diffractograms of the obtained films are shown in Fig.1. The analysis of the diffractograms shows that the structure of the obtained films corresponds to the hexagonal crystal structure of GaN wurtzite.

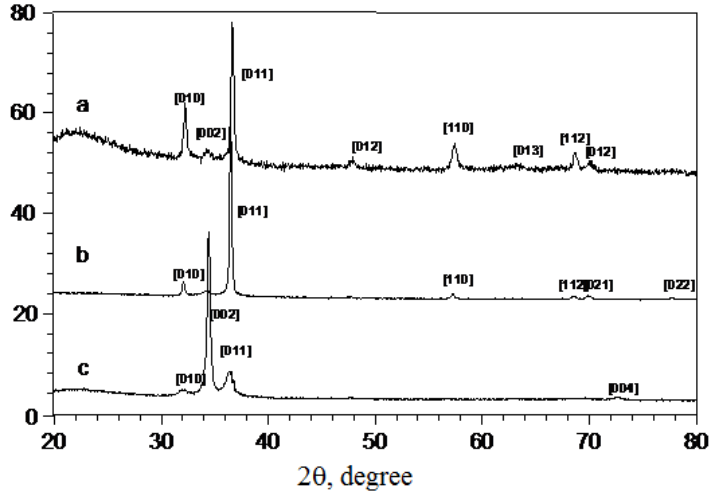


Fig. 1. X-ray diffractograms of GaN thin films deposited on a sapphire substrate with a $MgAl_2O_4$ sublayer (a), on a pure sapphire substrate (b), and on a quartz substrate with a ZnO sublayer (c).

The results obtained indicate that the films are characterized by a polycrystalline structure. At the same time, when deposited on sapphire substrates, the predominant orientation of the films is observed in the (011) plane, and when deposited on quartz substrates, in the (002) plane. The use of a sublayer does not change this pattern but leads to a slight increase in the orientation of the films in other planes. For example, the use of a $MgAl_2O_4$ sublayer on sapphire substrates leads to a slight increase in the orientation of GaN films in the (010), (110), and (112) planes (Fig. 1, *a* and *b*).

Conduction currents in the temperature range of 100-450 K were measured on an automated setup. An electric voltage of 10-100 V was applied to two-point contacts with a diameter of 1 mm, which were applied at a distance of 1 mm. When measuring the current flowing in GaN thin films, the main requirement is the use of ohmic, non-rectifying contacts that do not create additional barriers at the interface. The ohmic contact to the studied films is created by materials that, when displaced directly, provide electron injection into the film and have an output work of ~4.5 eV. The polycrystalline carbon (Aquadag) we use meets these requirements and has been used in numerous publications to study diamond, garnet, spinel, and other high-resistivity samples [7-9].

RESULTS AND DISCUSSION

The studies show that GaN thin films deposited on sapphire substrates, depending on the preparation conditions, have a rather high resistivity from 10^5 to 10^{10} Ohm×cm. The temperature dependence of the electrical conductivity for such films in the coordinates ln*I*= f(1000/T) is shown in Fig. 2 (curve 1). As can be seen from the dependence, the experimental results obtained in the temperature range of 100..450 K are well approximated by a linear dependence. This situation indicates the activation nature of the electrical conductivity of these films and allows us to determine the energy of thermal activation of electrical conductivity, which is E_T = 0.34 eV, from the slope of the line.

It should be noted that the decrease in resistivity in gallium nitride films deposited on a ZnO sublayer may be due to the interaction between the sublayer material and the film material. In particular, on the one hand, the possible diffusion of excess Ga^{3+} from the film can lead to an increase in the conductivity of ZnO sublayers, and on the other hand, the

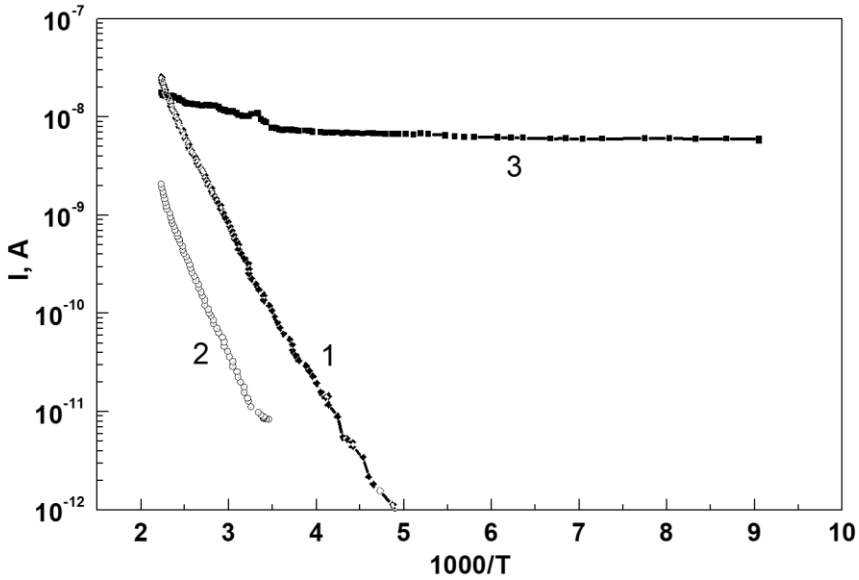


Fig. 2. Temperature dependence of the electrical conductivity of GaN thin films deposited on a sapphire substrate without a buffer sublayer (1) and with a MgAl2O4 sublayer (2), films on quartz substrates with a ZnO sublayer (3).

diffusion of Zn^{2+} into GaN can lead to an increase in conductivity due to acceptor centers in gallium nitride. It is known that GaN can have both impurity electronic and hole conductivity [10-12]. The donors in GaN films can be substitutional atoms Si_{Ga}, Ge_{Ga}, O_N, interstitial gallium Ga_i, and nitrogen vacancies V_N. Such donors create low energy levels in the band gap near the bottom of the conduction band.

For a more detailed study of the electrical conductivity of GaN thin films, the effect of the technological conditions for obtaining the films, in particular the pressure of the working gas N_2 and the temperature of the substrate, was investigated. The dependence of the specific electrical conductivity of GaN thin films on sapphire substrates on the partial pressure of N_2 in the working chamber during the sputtering of the films is shown in Fig. 3.

As can be seen from Fig. 3, with an increase in the partial pressure of nitrogen from 0.006 to 0.03 Torr, the value of the specific electrical conductivity of GaN films decreases by several orders of magnitude from 10^{-5} to 10^{-9} $Ohm^{-1}cm^{-1}$. Such a significant dependence of the electrical conductivity of the obtained films on the partial pressure of N_2 indicates that the most dominant charge carriers in gallium nitride thin films under RF ion-plasma sputtering are nitrogen vacancies, which play the role of electron donors. A decrease in the partial pressure of nitrogen in the sputtering atmosphere leads to an increase in nitrogen vacancies in the films and, accordingly, to an increase in the donor electrical conductivity.

It has been found that the temperature of the substrate largely determines the electrical conductivity of gallium nitride thin films during film deposition. The results of measuring the specific electrical conductivity at 300 K for GaN films obtained at different substrate temperatures are shown in Fig. 4. As can be seen from the figure, an increase in the substrate temperature from 500°C to 630°C almost linearly leads to a significant increase in the specific electrical conductivity from 10^{-9} to 10^{-5} $Ohm^{-1}cm^{-1}$. Such a significant increase in the specific electrical conductivity with increasing substrate temperature indicates the formation of electrically active point defects.

The obtained dependences of the specific electrical conductivity σ of the obtained gallium nitride thin films on the partial pressure of nitrogen and the temperature of the substrate can be explained if we consider that the most likely donor-type defects in these

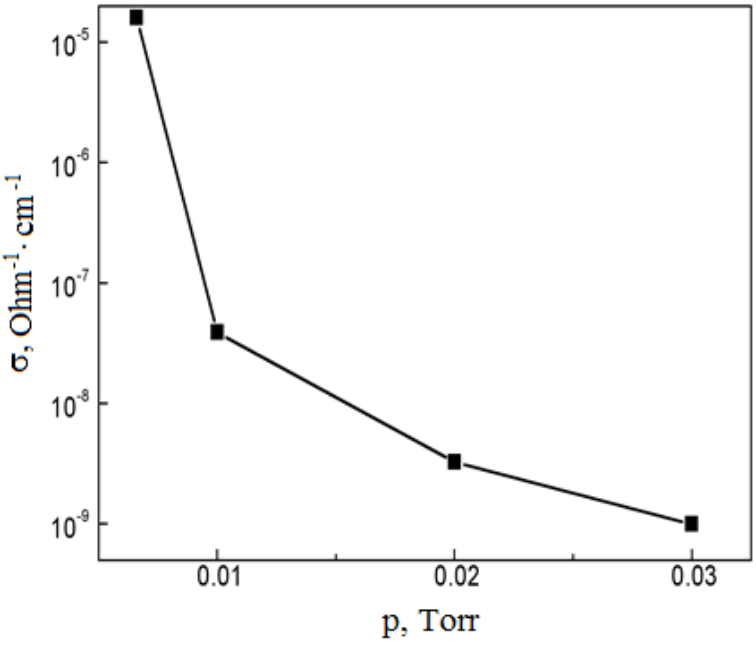


Fig. 3. Dependence of the specific electrical conductivity (at 300 K) of GaN thin films on sapphire substrates on the partial pressure of N2 during film sputtering (substrate temperature at 800 K).

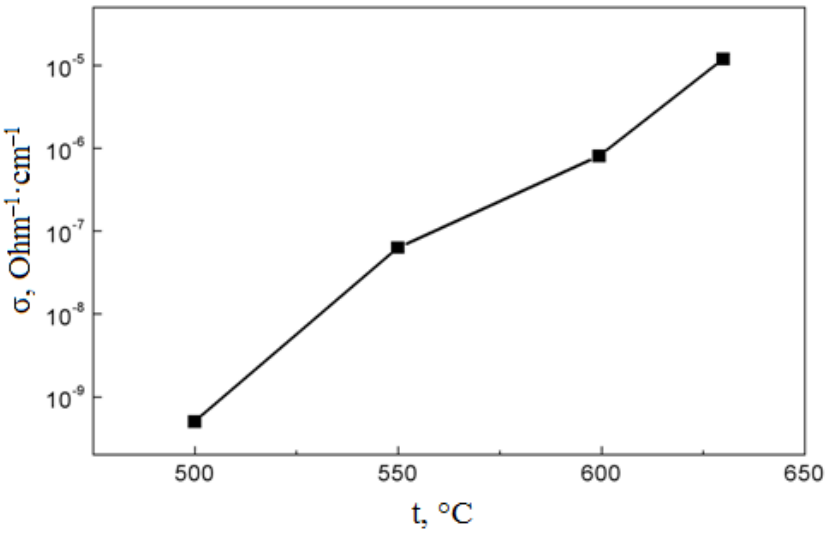


Fig. 4. Dependence of the specific electrical conductivity of GaN thin films on sapphire substrates at 300 K on the temperature of the substrate during sputtering.

films are nitrogen vacancies V_N. According to [13], epitaxial GaN films have an electron concentration that can vary from 10^{17} to 10^{20} cm^{-3} depending on the preparation conditions. The experimental values of the energy levels of donor centers in pure GaN films and crystals are conditionally divided into 4 ranges. This is the range of very low energies of 0.01..0.04 eV, intermediate energies of about 0.1 eV, and average energies of 0.2..0.4 eV [14-17], as well as in the region of 0.33..0.39 eV for high-resistivity gallium nitride samples [17]. Such differences in the values of the activation energy of conductivity in gallium nitride are associated with the existence of several types of donor defects and the interaction of point defects with the formation of defect complexes. Our experimental values for the thermal activation energy E_T = 0.34 eV, for GaN thin films on sapphire substrates are consistent with the literature data [17] for the energy depth of donor levels in high-resistivity gallium nitride films. For GaN films on quartz substrates with a ZnO sublayer, the determined values of E_T = 0.004 eV and E_T = 0.05 eV fall into the range of very low energies and are characteristic of low-impedance gallium nitride samples.

In [14, 18, 19], theoretical calculations of the formation energies of the most probable point defects in GaN were performed. According to these calculations, when thin films are deposited in a low-pressure nitrogen atmosphere, there are two dominant types of defects: nitrogen vacancies V_N, and inter-nodal gallium Ga_i. It is shown that the energy of nitrogen vacancy formation is several times lower than the energy of gallium vacancy formation and other point defects. Therefore, given that the energy of formation of nitrogen vacancies in gallium nitride is much lower than the energy of formation of other point defects, the most likely donor defects in this semiconductor should be considered anionic vacancies V_N.

CONCLUSION

The studies have shown that the RF ion-plasma sputtering of GaN thin films in a nitrogen atmosphere produces films whose electrical conductivity is determined by the type of substrate, the presence of a buffer sublayer, the pressure of the working gas, and the temperature of the substrate during sputtering. The thermal activation energy of electrical conductivity of high-resistance films is 0.34 eV. For low-impedance gallium nitride films, this value ranges from 0.004 eV to 0.05 eV, depending on the temperature range. The obtained dependences of the change in the electrical conductivity of GaN thin films on the partial pressure of the working gas nitrogen and the temperature of the substrate during

sputtering show that the most likely donor-type defects in gallium nitride are nitrogen vacancies V_N.

COMPLIANCE WITH ETHICAL STANDARDS

The authors declare that they have no competing interests.

AUTHOR CONTRIBUTIONS

Conceptualization, [O.B., I.Ku., M.P., I.M., I.Ko., Zh. T.]; methodology, [O.B., I.Ku.]; validation, [M.P., I.M.]; investigation, [I.Ku., M.P., I.M., I.Ko.]; writing – original draft preparation, [M.P., I.M., I.Ko., Zh.T.]; writing – review and editing, [O.B., I.Ku.]; visualization, [O.B., I.Ku., M.P., I.M., I.Ko., Zh.T.].

All authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] Bruckbauer, J., Cios, G., Sarua, A., Feng, P., Wang, T., Hourahine, B., Winkelmann, A., Trager-Cowan, C., & Martin, R.W. (2025). Strain and luminescence properties of hexagonal hillocks in N-polar GaN. *Journal of Applied Physics,* 137, 135705. https://doi.org/10.1063/5.0259840

[2] Khan, M.A.H., & Rao, M.V. (2020). Gallium Nitride (GaN) Nanostructures and Their Gas Sensing Properties: A Review. *Sensors*, 20, 3889. https://doi.org/10.3390/s20143889

[3] Pearton, S.J., Zolper, J.C., Shul, R.J., & Ren, F. (1999). GaN: Processing, defects, and devices. *Journal of Applied Physics,* 86, 1–78. https://doi.org/10.1063/1.371145

[4] Tian, J., Lai, C., Feng, G., Banerjee, D., Li, W., & Kar, N.C. (2020). Review of recent progresses on gallium nitride transistor in power conversion application. *International Journal of Sustainable Energy*, 39 (1), 88–100. https://doi.org/10.1080/14786451.2019.1657866

[5] Lidow, A., De Rooij, M., Strydom, J., Reusch, D., & Glaser, J. (2020). GaN Transistors for Efficient Power Conversion. Wiley.

[6] Wasa, R., Kitabatake, M., & Adachi, H. (2004). Thin Film Materials Technology: Sputtering of Compound Materials. William Andrew Publishing.

[7] Sinkler, W., Marks, L.D., Edwards, D.D., Mason, T.O., Poeppelmeier, K.R., Hu, Z., & Jorgensen J.D. (1998). Determination of Oxygen Atomic Positions in a Ga–In–Sn–O Ceramic Using Direct Methods and Electron Diffraction. *Journal of Solid State Chemistry*, 136 (1), 145–149. https://doi.org/10.1006/jssc.1998.7804

[8] Vasyltsiv, V.I., Rym, Ya.I., & Zakharko, Ya.M. (1996). Optical absorption and photoconductivity at the band edge of β-$Ga_{2-x}In_xO_3$. *Physica Status Solidi* (*b*), 195 (2), 653–658. https://doi.org/10.1002/pssb.2221950232

[9] Bordun, O. M., Kukharskyy, I. Yo., Medvid, I.I., Maksymchuk, D. M., Ivashchyshyn, F. O., Całus, D., & Leonov, D. S. (2022). Electrical Conductivity of Pure and Cr^{3+} -Doped β-Ga_2O_3 Thin Films. *Nanosistemi, Nanomateriali, Nanotehnologii*, 20(2), 321–329.

[10] Monish, M., Mohan, S., Sutar, D.S., & Major, S.S. (2020). Gallium nitride films of high n-type conductivity grown by reactive sputtering. *Semiconductor Scince Technology*, 35 (4), 045011. https://doi.org/10.1088/1361-6641/ab73ec

[11] Loretz, P., Tschirky, T., Isa, F., Patscheider, J., Trottmann, M., Wichser, A., Pedrini, J., Bonera, E., Pezzoli, F., & Jaeger, D. (2022). Conductive n-type gallium nitride thin films prepared by sputter deposition. *Journal of Vacuum Scince & Technology A*, 40, 042703. https://doi.org/10.1116/6.0001623

[12] Quan, Y., Yue, S.-Y., & Liao, B. (2021). Electric field effect on the thermal conductivity of wurtzite GaN. *Applied Physics Letters*, 118 (16), 162110. https://doi.org/10.1063/5.0047372

[13] As, D.J., Schikora, D., Greiner, A., Lübbers, M., Mimkes, J., & Lischka, K. (1996). p- and n-type cubic GaN epilayers on GaAs. *Physical Review B*, 54, R11118 – R11121. https://doi.org/10.1103/PhysRevB.54.R11118
[14] Perlin, P., Suski, T., Teisseyre, H., Leszczynski, M., Grzegory, I., Jun, J., Porowski, S., Bogusławski, P., Bernholc, J., Chervin, J.C., Polian, A., & Moustakas, T.D. (1995). Towards the Identification of the Dominant Donor in GaN. *Physical Review Letters*, 75, 296. https://doi.org/10.1103/PhysRevLett.75.296
[15] Lyons, J.L., Wickramaratne, D., & Van de Walle, C.G. (2021). A first-principles understanding of point defects and impurities in GaN. *Journal of Applied Physics*, 129, 111101. https://doi.org/10.1063/5.0041506
[16] Kaschner, A., Kaczmarczyk, G., Hoffmann, A., Thomsen, C., Birkle, U., Einfeldt, S., & Hommel, D. (1999). Defect Complexes in Highly Mg-Doped GaN Studied by Raman Spectroscopy. *Physica Status Solidi* (*b*), 216 (1), 551–555. https://doi.org/10.1002/(SICI)1521-3951(199911)216:1<551::AID-PSSB551>3.0.CO;2-S
[17] Strauf, S., Michler, P., Gutowski, J., Birkle, U., Fehrer, M., Einfeldt, S., & Hommel, D. (1999). Optical Spectroscopy of Mg- and C-Related Donor and Acceptor Levels in GaN Grown by MBE. *Physica Status Solidi* (*b*), 216 (1), 557–560. https://doi.org/10.1002/(SICI)1521-3951(199911)216:1<557::AID-PSSB557>3.0.CO;2-4
[18] Neugebauer, J., & Van de Walle, C.G. (1994). Atomic geometry and electronic structure of native defects in GaN. *Physical Review B*, 50, 8067. https://doi.org/10.1103/PhysRevB.50.8067
[19] Lyons, J.L., & Van de Walle, C.G. (2017). Computationally predicted energies and properties of defects in GaN. *npj Computational Materials*, 3, 12. https://doi.org/10.1038/s41524-017-0014-2

ВПЛИВ УМОВ ОДЕРЖАННЯ НА ЕЛЕКТРОПРОВІДНІСТЬ ТОНКИХ ПЛІВОК GaN

Олег Бордун**, Ігор Кухарський, Марія Процак, Іванна Медвідь, Ірина Кофлюк, Жанетта Цаповська
Львівський національний університет імені Івана Франка,
вул. Драгоманова 50, 79005 м. Львів, Україна

АНОТАЦІЯ

Вступ. Нітрид галію (GaN) — перспективний матеріал для створення світлодіодів, лазерів, детекторів ультрафіолету й високочастотних транзисторів. Властивості GaN суттєво залежать від типу дефектів та параметрів одержання тонких плівок. Зокрема, вплив мають підкладки, буферні шари, температура напилення та парціальний тиск робочого газу. Це обумовило дослідження електропровідності тонких плівок GaN, отриманих методом високочастотного (ВЧ) іонно-плазмового розпилення за різних технологічних умов.

Матеріали та методи. Плівки GaN товщиною 0,3–1,0 мкм осаджували на сапфірові та кварцові підкладки, з/без буферних шарів $MgAl_2O_4$ або ZnO. ВЧ розпилення проводилось в атмосфері N_2 при тиску $5\cdot10^{-3}$–$5\cdot10^{-2}$ Торр та температурах 400–650 °C. Структуру аналізували за допомогою Х-променевої дифракції, електропровідність визначали в температурному діапазоні 100–450 К методом двоточкових контактів з використанням омічних контактів із вуглецю.

Результати. Встановлено, що GaN-плівки мають полікристалічну структуру з різною орієнтацією, залежно від типу підкладки та наявності буферного шару. Питомий опір змінюється від 10^5 до 10^{10} Ом·см. Високоомні зразки характеризуються енергією термічної активації 0,34 еВ та встановлено донорний тип провідності у плівках. У плівках на кварцових підкладках із підшаром ZnO зафіксовано менші значення енергії активації (0,004..0,05 еВ), що свідчить про інший механізм провідності. Провідність GaN зростає при збільшенні температури підкладки і зниженні парціального тиску N_2 в розпилювальній атмосфері, що приводить до утворення вакансій азоту — основних донорних дефектів.

Висновки. Електропровідність тонких плівок GaN значною мірою визначається умовами їх одержання. Найбільший вплив мають тип підкладки, буферні шари, температура напилення та тиск газу в розпилювальній атмосфері. Отримані залежності зміни величини електропровідності тонких плівок GaN залежно від парціального тиску робочого газу азоту та температури підкладки при напиленні показують, що найбільш імовірними дефектами донорного типу в нітриді галію є вакансії азоту V_N.

Ключові слова: тонкі плівки, нітрид галію, електропровідність, дефекти донорного типу.

Received / Одержано	Revised / Доопрацьовано	Accepted / Прийнято	Published / Опубліковано
25 April, 2025	26 May, 2025	02 June, 2025	26 June, 2025

Electronics and Information Technologies, 2025, 30, 163–176
http://publications.lnu.edu.ua/collections/index.php/electronics/index

UDC: 538.975, 539.2, 621.382

MAGNETO- AND PHOTO-CONTROLLED SUPRAMOLECULAR CLATRATES BASED ON GALLIUM SELENIDE

***Vitalii Maksymych*[1]*, Oleksii Korchev*[2]*, Nadiia Pokladok*[2]*, Roman Shvets*[2]*, Oleh Bordun*[1]*, Fedir Ivashchyshyn*[3]**

[1]*Faculty of Electronics and Computer Technologies, Ivan Franko National University of Lviv, 50 Dragomanov Str., 79005 Lviv, Ukraine*

[2]*Institute of Applied Mathematics and Fundamental Sciences, Lviv Polytechnic National University, 12 Stepan Bandera Str., 79013 Lviv, Ukraine*

[3]*Faculty of Electrical Engineering, Czestochowa University of Technology, ul. J.H. Dąbrowskiego 69, 42-201 Częstochowa, Poland*

Maksymych, V., Korchev, O., Pokladok, N. et al. (2025). Magneto- and Photo-Controlled Supramolecular Clatrates Based on Gallium Selenide. *Electronics and Information Technologies*, 30, 163–176. https://doi.org/10.30970/eli.30.13

ABSTRACT

Background. The development of modern nanoelectronics and sensor technologies necessitates the creation of new functional materials with tunable electrical, optical, and magnetic properties. One promising approach involves the formation of supramolecular clathrates based on layered A^3B^6 group semiconductors. This paper presents the results of electrophysical analysis of GaSe-based clathrate systems intercalated with a β-cyclodextrin<ferrocene> (β-CD<FC>) complex, synthesized both under normal conditions and in a constant magnetic field.

Materials and Methods. A GaSe single crystal was used as the host matrix. The β-CD<FC> guest complex was intercalated under two different conditions: normal conditions and a constant magnetic field with a strength of 220 kA/m. Electrophysical properties were studied using impedance spectroscopy over a frequency range of 10^{-3}–10^6 Hz and thermally stimulated discharge analysis.

Results and Discussion. Intercalation of β-CD<FC> into the expanded GaSe structure results in increased resistance in the low-frequency range and a non-monotonic behaviour of ReZ(ω). In samples synthesized in a magnetic field, a decrease in ReZ(ω) is observed, indicating a change in the electronic character of the guest complex from acceptor to donor type. Thermally stimulated discharge spectra reveal a transition from a quasi-continuous energy level distribution to a mini-zone structure, with homocharge relaxation dominating in the clathrate phase. Furthermore, samples synthesized in a magnetic field exhibit a fivefold increase in the magnetoresistive effect and a twofold enhancement in photosensitivity.

Conclusion. The study demonstrates that intercalation of the supramolecular β-CD<FC> complex into the expanded GaSe matrix enables targeted modification of the impurity structure, resulting in changes in electrical conductivity and sensory behaviour of the clathrate. It was established that synthesis conditions, particularly the presence of a constant magnetic field, significantly influence the electronic nature of the guest component, the type of charge carriers, trap level parameters, and the manifestation of quantum effects during charge transport.

Keywords: Intercalation, impedance spectroscopy, magnetoresistive effect, photoresistive effect, supramolecular complex, hierarchical architecture.

INTRODUCTION

The rapid advancement of modern electronics and sensor technologies necessitates the development of new functional materials with tunable electrical and physical characteristics. Such materials can be designed based on nanostructured systems capable of exhibiting novel physical phenomena due to their internal architecture, engineered at the atomic and molecular levels. Particularly noteworthy in this context are hybrid nanostructured systems that demonstrate complex sensitivity to external fields, thereby opening new avenues for their integration into multifunctional device components.

One promising direction in this field involves the formation of clathrates–supramolecular compounds in which guest molecules or complexes are incorporated into the cavities of a host matrix [1–3]. In particular, the intercalation of functionalized organic complexes into layered semiconducting materials offers a versatile approach to tailoring the electronic, magnetic, and optical properties of the resulting systems.

Intercalation into A^3B^6-type layered semiconductor crystals, such as gallium selenide (GaSe), represents an effective method for modifying their electronic structure, impurity spectrum, and physicochemical characteristics. Owing to the presence of weak interlayer van der Waals interactions, these crystals can accommodate guest species—particles, molecules, or complexes—within their interlayer spaces without disrupting the integrity of the internal crystal lattice [4, 5]. This enables the formation of stable intercalated systems that preserve their layered morphology while acquiring novel functional properties, including enhanced electrical conductivity, magnetic responsiveness, and photoelectric behaviour.

Of particular scientific interest is the formation of clathrates with a hierarchical architecture of the subhost<host<guest>> type. In this regard, the intercalation of supramolecular complexes such as β-cyclodextrin with ferrocene (β-CD) appears especially promising, as these components can serve not only as carriers of functional properties (e.g., photo- or magneto-sensitivity) but also as modulators of electronic state ordering due to their specific geometry and inherent self-assembly capabilities.

This work presents the results of a study on clathrates formed by two approaches: intercalation of the β-CD complex into the GaSe matrix under normal conditions and intercalation performed in the presence of a constant magnetic field.

MATERIALS AND METHODS

In the experiments, a semiconductor single crystal of gallium selenide (GaSe) (Fig. 1) was used as the subhost material. Grown by the Bridgman–Stockbarger method, the single crystal exhibits a multilayer hexagonal structure and p-type conductivity. GaSe is particularly well-suited for intercalation owing to its wide band gap (2.02 eV, as determined from optical data), high anisotropy of electrical conductivity, and resistance to chemical modification. It is also characterised by the presence of so-called guest-accessible sites — regions of weak van der Waals forces oriented perpendicular to the crystallographic C-axis [5].

β-cyclodextrin ($C_{42}H_{70}O_{35}$) (β-CD) [6] was selected as the host. This compound is a cyclic oligosaccharide with a well-defined toroidal geometry (see Fig. 2), featuring a hydrophilic outer surface and a hydrophobic inner cavity. Such a structure enables the formation of inclusion complexes with hydrophobic guest molecules via weak non-covalent interactions, particularly van der Waals forces. Owing to its rigid geometry and propensity for forming ordered supramolecular assemblies, β-CD serves as a structural platform for the spatial organisation of guest species within the interlayer regions of host materials [7-9].

Ferrocene (FC) (Fig. 3) was used as the guest molecule. Ferrocene is an organometallic compound with well-characterised donor–acceptor and electrochemical properties and exhibits high stability in a wide range of chemical environments. Its molecular structure consists of an iron atom sandwiched between two cyclopentadienyl rings, forming a delocalized π-system that facilitates interaction with the electronic states of the sub-host material [10-12].

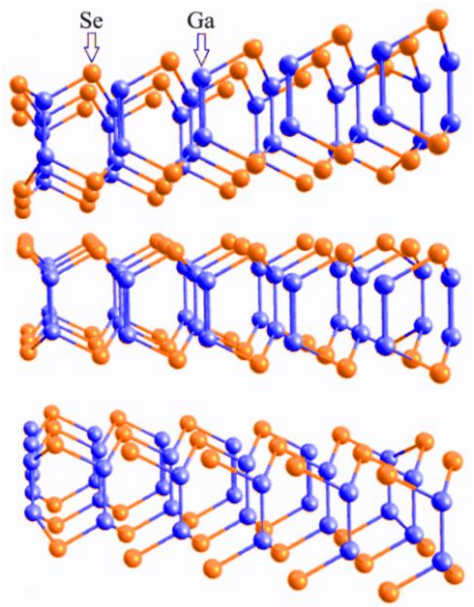


Fig. 1. Spatial representation of the GaSe structure.

Ferrocene (FC) (Fig. 3) was used as the guest molecule. Ferrocene is an organometallic compound with well-characterised donor–acceptor and electrochemical properties and exhibits high stability in a wide range of chemical environments. Its molecular structure consists of an iron atom sandwiched between two cyclopentadienyl rings, forming a delocalized π-system that facilitates interaction with the electronic states of the sub-host material [10-12].

Since ferrocene is insoluble in water, the supramolecular complex β-CD<FC> was prepared using a method described in [13]. A fine powder of ferrocene was added to an aqueous solution of β-CD at 60 °C under continuous stirring. The molar ratio of β-CD to ferrocene was maintained at 1:1.

The intercalation of the β-CD<FC> supramolecular complex into GaSe crystals involved several stages and has been described in detail in our previous works [14, 15]. First, sodium nitrite ($NaNO_2$) was introduced into the GaSe matrix until a 5-fold expansion was achieved. Following this, the $NaNO_2$ was washed out, and the crystals were dried in a vacuum oven at 100 °C for 24 hours.

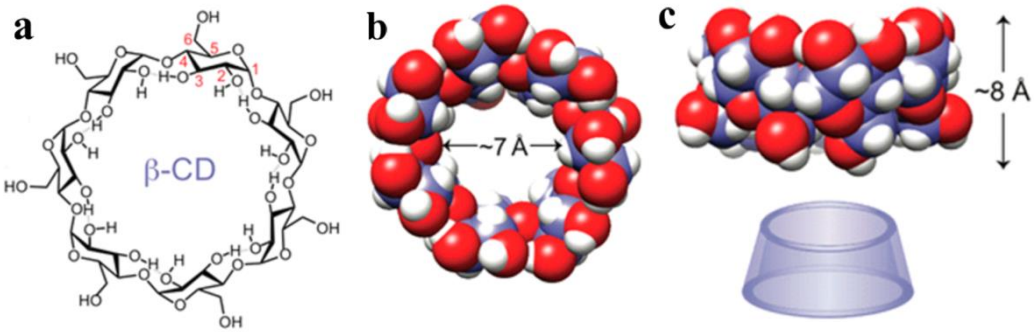


Fig. 2. Structure of β-cyclodextrin (β-CD): schematic representation of the molecule with atom numbering of the glucopyranose units (a); space-filling model showing the inner cavity (b); truncated cone-like shape of β-CD (c).

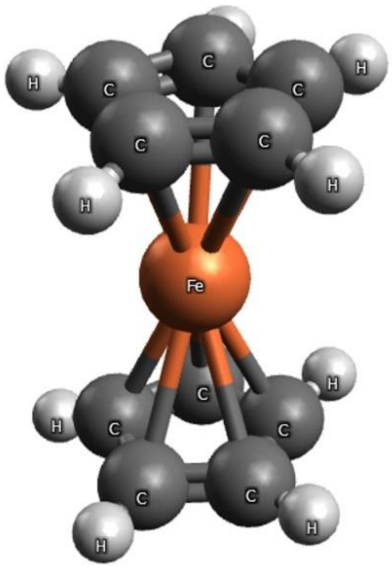


Fig. 3. Molecular structure of ferrocene.

The intercalation of β-CD<FC> was performed in two modes. In the first mode, the GaSe crystal was immersed in an aqueous solution containing β-CD and ferrocene under normal conditions. In the second mode, to control the spatial ordering of the magnetically sensitive guest during intercalation, the GaSe crystal was placed in the same solution while exposed to a constant magnetic field with a strength of 220 kA/m. During the intercalation of the supramolecular complex β-CD<FC>, the GaSe crystal was positioned so that the magnetic induction vector was oriented parallel to the crystallographic axis C, i.e., perpendicular to the plane of the layers. This orientation ensures maximum influence of the magnetic field on the ordering of magnetosensitive ferrocene molecules along the interlayer voids of the crystal. Upon completion, the intercalated GaSe<β-CD<FC>> clathrate samples were dried at 45 °C for 24 hours in a vacuum oven.

As a result, the obtained experimental sample represents a stage-wise ordered layered structure in which unexpanded GaSe layers alternate with expanded layers containing the intercalated β-CD<FC> supramolecular complex.

The studies using impedance spectroscopy were carried out along the crystallographic axis C in a $10^{-3}..10^{6}$ Hz frequency range using the AUTOLAB measurement system (ECO CHEMIE, the Netherlands), equipped with FRA-2 and GPES software packages. The uncertain points were removed using the Dirichlet filter [16]. The impedance measurements were made under normal conditions, in a constant magnetic field of 220 kA/m lighted by a 65W solar simulator. The external fields were applied along the impedance spectra measurements. This geometry of measurements was chosen in order to dismiss the Lorentz force.

The studies were also carried out using the thermally stimulated discharge in the temperature range of –25..+70°C with a constant heating rate of 5 degrees/min.

Based on the obtained impedance spectroscopy data according to the Geballe-Pollack theory [17], the following parameters of the impurity energy spectrum were calculated: density of states at the Fermi level N_F, hopping radius R, spread of traps near the Fermi level J and real density of deep traps N_t.

RESULTS AND DISCUSSION

To investigate the influence of the guest molecule on the electrical conductivity of the GaSe<β-CD<FC>> clathrate, impedance spectroscopy was employed. The analysis

begins with a comparison of the real part of the complex impedance (ReZ) for the pristine GaSe single crystal, the 5-fold expanded GaSe matrix, and the GaSe<β-CD<FC>> clathrate synthesized under both normal conditions and in a constant magnetic field with an intensity of 220 kA/m. The frequency dependencies of ReZ are shown in Fig. 4.

The unmodified GaSe single crystal (curve 1, Fig. 4) exhibits frequency-independent behaviour of ReZ in the $10^{-2}..10^{2}$ Hz range, indicating conductivity governed by equilibrium charge carriers at the given temperature. This conductivity is of the activation type and can be described as follows:

$$\sigma_0 = en\mu, \tag{1}$$

where e is the elementary charge, n is the carrier concentration, and μ is their mobility.

A further increase in frequency ($\omega > 10^2$ Hz) leads to a monotonic decrease in the real part of the impedance (ReZ) of the GaSe single crystal due to the dominance of hopping conductivity. This conductivity arises from the hopping of nonequilibrium charge carriers between localised states near the Fermi level. In this process, the carriers hop from one centre localised in the forbidden band of the semiconductor to another, accompanied by the emission or absorption of a phonon. This type of conductivity is of the activation type, meaning that:

$$\sigma \sim \exp\left[-\frac{\Delta W}{kT}\right]. \tag{2}$$

In the case of alternating current, the hopping conductivity becomes frequency-dependent:

$$\sigma_j(\omega) \sim \omega^n, \tag{3}$$

where n the exponent typically ranges from 0.64 to 1.0.

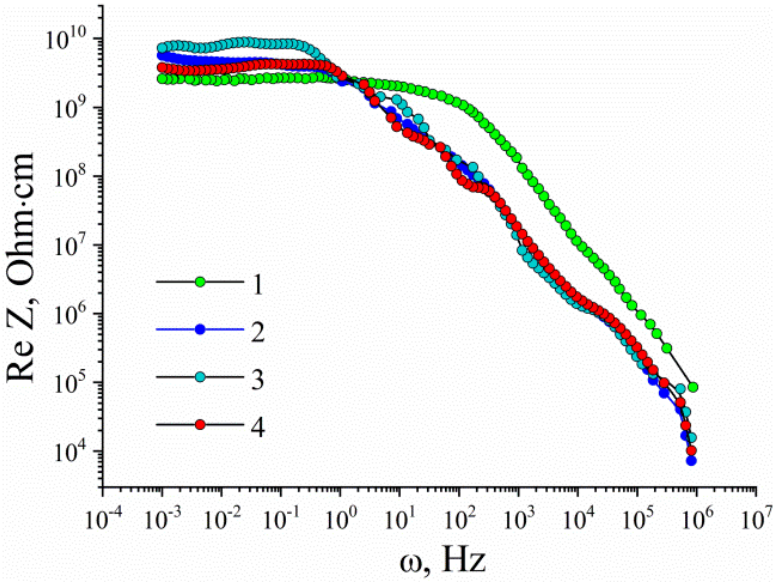


Fig. 4. Real part of complex impedance measured for the initial unexpanded GaSe crystal (1), 5-fold expanded GaSe matrix (2), and GaSe<β-CD<FC>> clathrate synthesized under normal conditions (3) and in a constant magnetic field (4).

As a result, the total conductivity of the GaSe single crystal can be expressed as a sum of equilibrium and hopping contributions:

$$\sigma'(\omega) = en\mu + A\omega^n. \quad (4)$$

Following the 5-fold expansion of the GaSe single crystal (curve 2, Fig. 4), significant changes in ReZ(ω) are observed. The resistance increases by more than a factor of two, attributed to a reduction in the concentration of equilibrium charge carriers due to the formation of additional recombination centres caused by the expansion of the crystal layers. The low-frequency branch of ReZ(ω) loses its frequency independence and begins to decrease monotonically within the 10^{-2}–0.6 Hz range, followed by a sharp and non-monotonic decline with increasing frequency. This deviation from the behaviour predicted by Eq. (3) suggests that, in addition to band and hopping transport mechanisms, charge transfer also involves capture–retention–release processes associated with quantum wells, which are most likely formed in the regions of expanded layers.

Incorporation of the β-CD<FC> supramolecular complex into the interlayer space of the 5-fold expanded GaSe matrix under normal conditions (curve 3, Fig. 4) leads to an increase in ReZ(ω) in the low-frequency range and a further enhancement of the non-monotonic behaviour. This indicates the formation of additional quantum wells due to the presence of the guest component, which likely exhibits an acceptor-type nature.

When the same supramolecular complex is introduced under a constant magnetic field (curve 4, Fig. 4), ReZ(ω) decreases in the low-frequency range, although the non-monotonic behaviour remains pronounced. This also indicates the formation of additional quantum wells, but in this case, the guest molecule likely demonstrates a donor-type character. Given the opposite trends in the low-frequency behaviour of ReZ(ω) depending on the synthesis conditions, we can assume that the nature of the guest molecule differs: acceptor-like under normal conditions and donor-like under the influence of a magnetic field.

For a more detailed interpretation of the ReZ(ω) variations depending on the synthesis conditions, Fig. 5 presents the ratio of the real part of the complex impedance of the clathrate (ReZ_{Clat}) to that of the 5-fold expanded GaSe matrix ($ReZ_{Exp.mat}$). As can be seen, depending on the synthesis conditions, the $ReZ_{Clat}/ReZ_{Exp.mat}$ (ω) dependence exhibits

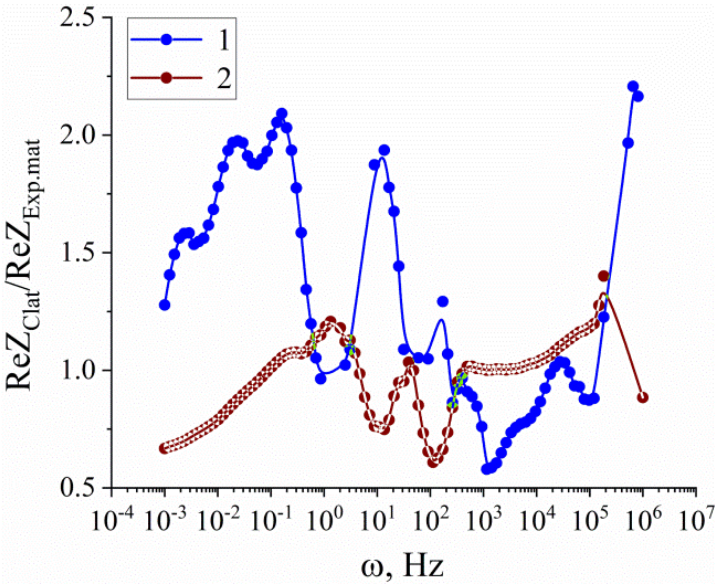


Fig. 5. Change in the real part of the complex impedance for GaSe<β-CD<FC>> clathrate synthesized under normal conditions (1) and in a constant magnetic field (2).

opposite frequency trends, most likely due to different types of conductivity exhibited by the guest component.

To elucidate the structure of impurity energy spectra in the initial GaSe single crystal, the 5-fold expanded GaSe matrix, and the GaSe<β-CD<FC>> clathrates synthesized under normal conditions and in a constant magnetic field, thermally stimulated discharge current spectra were measured, as shown in Fig. 6. For the initial GaSe single crystal, thermally stimulated discharge current signals begin to appear at 283 K. The spectrum initially displays a distinctly miniband-like character, which gradually transitions to a quasi-continuous form with increasing temperature. It is important to note that heterocharge relaxation is observed in this case. For the 5-fold expanded GaSe matrix, the TSDC spectrum exhibits a strongly pronounced miniband-like structure (Fig. 6), with a higher density of states concentrated in the low-temperature region—where the initial GaSe crystal had virtually no detectable thermally stimulated discharge current response. Homocharge relaxation is predominant in this case. The crystal expansion clearly leads to the formation of trap centres localised in narrow bands within the forbidden gap. This result supports the ReZ(ω) behaviour previously observed for the 5-fold expanded GaSe matrix (Fig. 4).

In the case of the GaSe<β-CD<FC>> clathrate, regardless of the synthesis conditions, the thermally stimulated discharge current spectrum maintains a strongly expressed miniband-like character throughout the entire studied temperature range, with homocharge relaxation remaining dominant. The observed differences depending on the synthesis method lie in the specific energy distribution of impurity levels, which ultimately determine the conductivity characteristics of the clathrates.

Using the theory of hopping conductivity proposed by M. Pollak and T.H. Geballe [17], which takes into account the charge carriers' hoppings between states localised in space

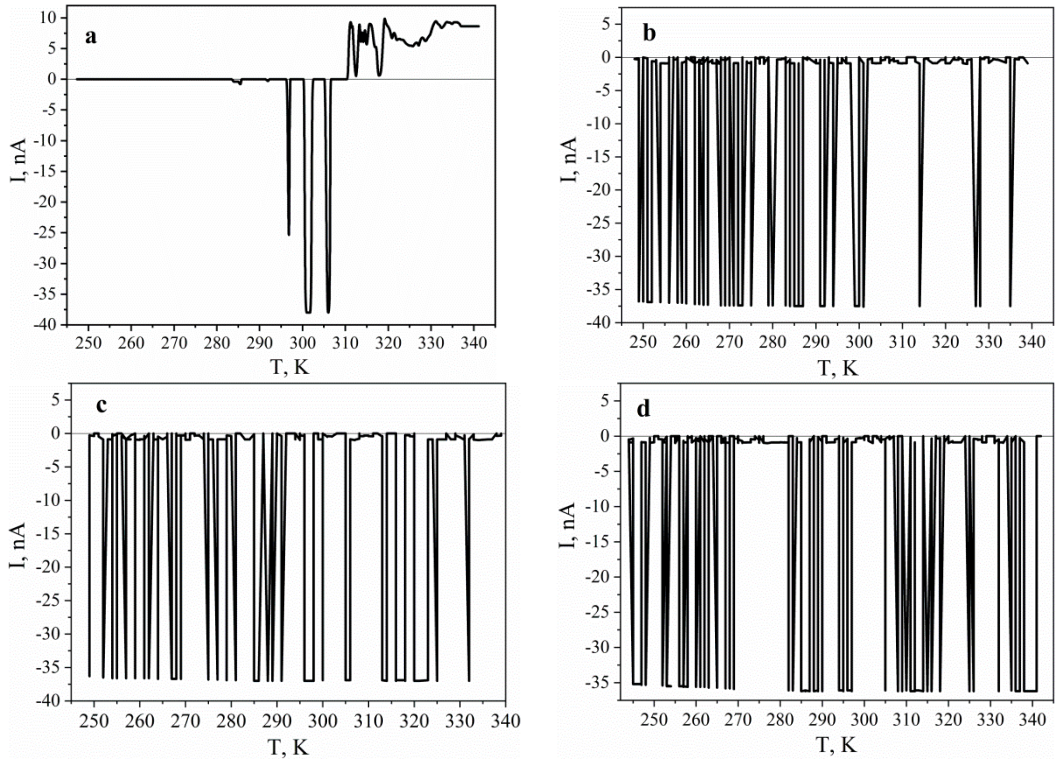


Fig. 6. Thermally stimulated discharge currents spectra for the initial unexpanded GaSe crystal (a), 5-fold expanded GaSe matrix (b), and GaSe<β-CD<FC>> clathrate synthesized under normal conditions (c) and in a constant magnetic field (d).

due to their interaction with phonons. The result is the following expression for the real part of the conductivity:

$$\sigma(\omega) = \frac{\pi}{96} e^2 k_B T \alpha^{-5} N_F^2 \omega \left[\ln\left(\frac{\nu_\Phi}{\omega}\right)\right]^4, \quad (5)$$

where e is the electron charge, N_F is the density of states on the Fermi level, α is the constant of the carrier localised wave function decrease, ν_Φ is the phonon frequency.

According to the theory of hopping conductivity on alternating current, the average hopping time τ of the carrier with phonon absorption or emission is determined by the expression:

$$\tau^{-1} = \nu_\Phi \exp(-2\alpha R), \quad (6)$$

where R is the hopping distance.

Experimentally, τ^{-1} is the average frequency at which the law $\omega^{0.8}$ is fulfilled. The average hopping distance R can be calculated by the formula after determining the value of τ^{-1} experimentally.

The scatter of trapped levels near the Fermi level J can be estimated by knowing the values of N_F and R from the relation

$$\frac{4}{3}\pi R^3 N_F \frac{1}{2} J = 1. \quad (7)$$

This, in turn, makes it possible to estimate the real density of deep traps N_t using the formula

$$N_t = N_F J. \quad (8)$$

The corresponding values obtained according to the presented calculations are given in Table 1. As can be seen, 5-fold expansion of the GaSe matrix leads to more than a twofold increase in the density of states at the Fermi level N_F, and a similar, but inverse, twofold decrease in the spread of trap levels J. The hopping distance R and the density of deep traps N_t remain unchanged. The introduction of the supramolecular complex β-CD<FC> into the interlayer space of the 5-fold expanded GaSe matrix results in a 1.82-fold increase in N_F and a 1.76-fold decrease in J. The hopping distance R slightly decreases, while the deep trap

Table 1. **Band spectrum parameters for the investigated samples measured at the normal conditions**

Structure	$N_F \cdot 10^{44}$, $J^{-1}m^{-1}$	$R \cdot 10^{-8}$, m	$J \cdot 10^{-22}$, J	$N_t \cdot 10^{22}$, m^{-3}
GaSe	0.34	2.55	8.42	2.89
5-fold expanded GaSe matrix	0.82	2.55	3.50	2.89
GaSe<β-CD<FC>> clathrate synthesized under normal conditions	1.50	2.52	1.99	3.00
GaSe<β-CD<FC>> clathrate synthesized in a constant magnetic field	0.75	2.80	2.89	2.17

density N_t increases somewhat. The formation of the GaSe<β-CD<FC>> deep trap density N_t increases somewhat. The formation of the GaSe<β-CD<FC>> clathrate in a constant magnetic field, compared to the 5-fold expanded GaSe matrix, leads to opposite effects: N_F decreases by about 10%, J decreases by 18%, N_t decreases by 25%, and R increases by 10%. Overall, there is a fairly good correlation between the aforementioned current transport mechanisms, the measured thermally stimulated discharge current spectra, and the theoretically calculated parameters of the impurity energy spectrum.

The next step was to investigate the influence of the guest component and the conditions of its formation on the magnitude of the magnetoresistive effect of the GaSe<β-CD<FC>> clathrate (Fig. 7). The magnetoresistive effect was calculated as the ratio of resistance in a constant magnetic field to resistance in its absence ($\delta_H=\rho_H/\rho_0$). As can be seen from Fig. 7, the synthesis conditions have a dramatic effect on the magnetoresistive effect, which differs in sign and magnitude. Thus, in the low-frequency range, a 5-fold increase in magnetosensitivity is observed for the GaSe<β-CD<FC>> clathrate synthesized in a constant magnetic field, while for the clathrate synthesized under normal conditions, a decrease of more than two times is observed. This indicates an increase in magnetosensitivity under conditions of clathrate synthesis in a constant magnetic field due to the ordering of the guest component.

Next, the influence of the guest component and its formation conditions on the magnitude of the photoresistive effect of the GaSe<β-CD<FC>> clathrate was investigated (Fig. 8). The photoresistive effect was calculated as the ratio of the resistance under illumination to the resistance in the dark ($\delta_L = \rho_L/\rho_D$). As seen in Fig. 8, the clathrate synthesis conditions result in an opposite frequency dependence of the photoresistive effect. In the low-frequency range, a twofold increase in photosensitivity is observed for the GaSe<β-CD<FC>> clathrate synthesized in a constant magnetic field, while a decrease is observed for the clathrate synthesized under normal conditions. This indicates enhanced photosensitivity when the clathrate is synthesized in a constant magnetic field, due to the formation of a specific impurity energy spectrum structure.

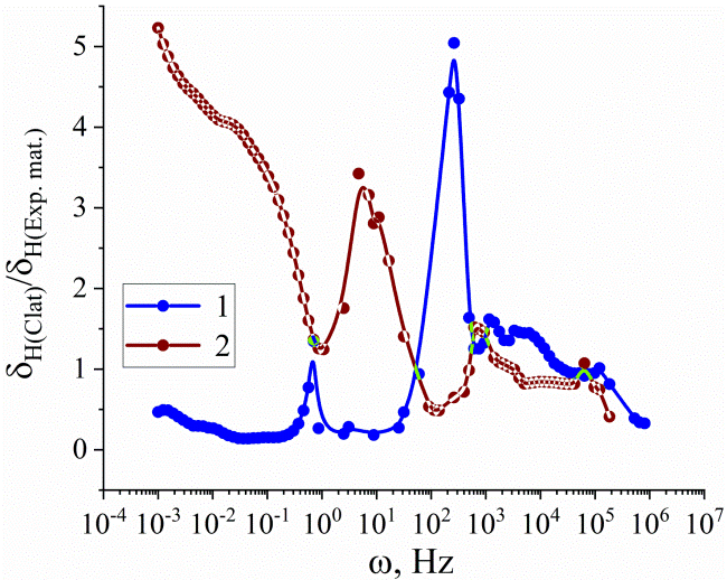


Fig. 7. Change in the magnetoresistive effect for the GaSe<β-CD<FC>> clathrate synthesized under normal conditions (1) and in a constant magnetic field (2).

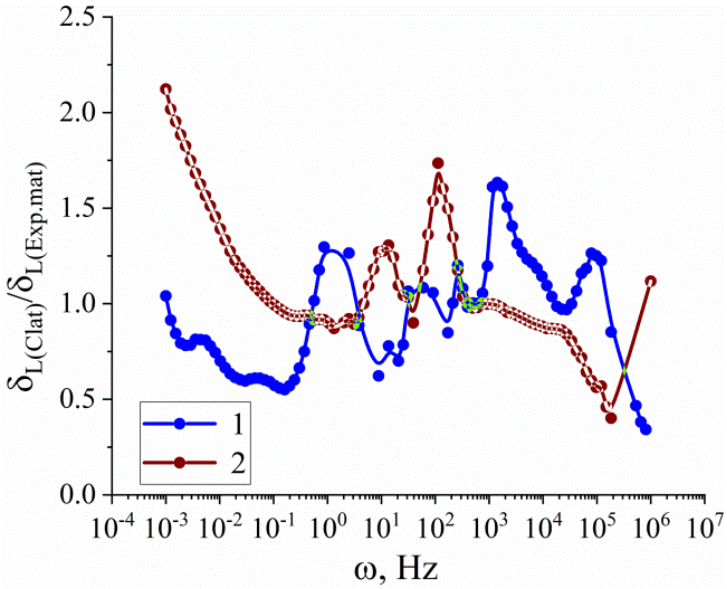


Fig. 8. Change in the photoresistive effect for GaSe<β-CD<FC>> clathrate synthesized under normal conditions (1) and in a constant magnetic field (2).

However, the most unexpected look is the I–V characteristics of the GaSe<β-CD<FC>> clathrate, regardless of the synthesis conditions, when measured under a constant magnetic field. (Fig. 9). While the I–V curves of the initial GaSe single crystal and the 5-fold expanded GaSe matrix exhibited a linear behaviour under all measurement conditions, the I–V characteristics of the GaSe<β-CD<FC>> clathrate synthesized under normal conditions and measured in a constant magnetic field demonstrate a somewhat unconventional appearance. In the presence of a magnetic field, the I–V curve takes on an oscillatory form.

This behaviour may be related to the following. Based on the architecture of the synthesized clathrate GaSe<β-CD<FC>> and the above research results, it is likely to assume that the formed nano-dispersed magnetic particles are an ensemble of quantum dots localised in the corresponding hierarchical structure. This means that the determining processes in charge transfer will be the tunnelling of current carriers and their limitation by a localised charge, like the Coulomb blockade [18]. Non-core current carriers will be localised in quantum wells, creating an energy barrier for the main ones. It is for such cases

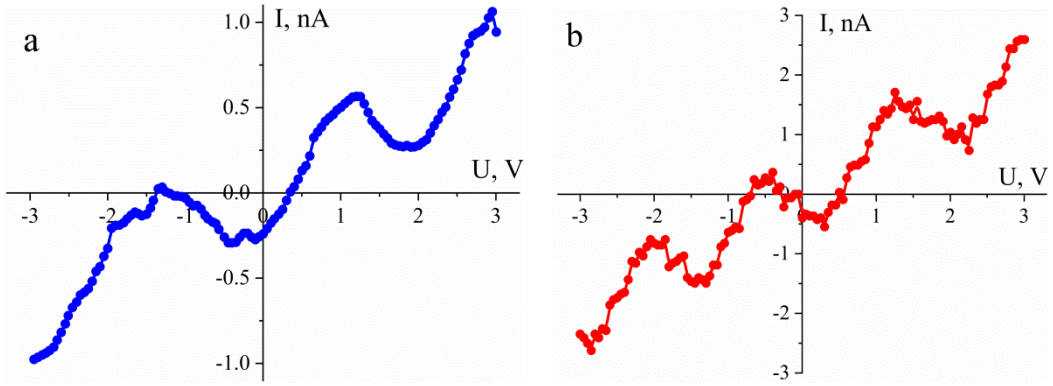


Fig. 9. I–V characteristics of GaSe<β-CD<FC>> clathrate synthesized under normal conditions (a) and in a constant magnetic field (b).

that the extraordinary behaviour of the dynamic volt-ampere characteristic is characteristic. The reason is that the “charging” and “discharging” of the 1D-barrier-0D nanocapacitor occur at different speeds: charging due to the inertial process of movement of delocalized media, and discharge due to their emission into the layer of traps with a lifetime τ. The competition of the processes of charging and discharging of these nanocapacitors in combination with the effect of the accumulation of a significant nonequilibrium charge determines the nature of the behaviour of the volt-ampere characteristic. It is clear that such effects will be the more significant, the greater the τ.

The qualitatively nonmonotonic nature of the volt-ampere characteristic can be understood based on a simple quantum mechanical model. Let us consider a one-dimensional crystal characterised by a band spectrum. The introduction of a molecule (atom or, in our case, a nanocluster) is accompanied by additional potential of the molecule. Depending on the specific content of the molecule, the potential can be both a barrier against the background of the crystal potential and a well.

Consider the case when the intercalated supramolecular complex generates additional localised potential (albeit rectangular) in the form of wells. In an ideal structure, where the geometric and energy characteristics of such wells are strictly maintained and the wells are positioned at fixed distances from each other, they are separated by the chain potential, which acts as a barrier between them. Electronic states in the wells are discrete. The voltage applied to such a structure can lead to the appearance of current by the mechanisms of jumping conductivity and/or resonant tunnelling. The attached magnetic field changes the position of the discrete levels. Each time such levels intersect with a fixed level, there is an abrupt change in carrier concentration, which is repeated with each subsequent intersection. The frequency of jumps is determined by the degree of discreteness of electronic states. Using the conclusions of the infinite deep well problem, namely

$$E_n = \frac{\pi^2 \hbar^2}{2ma^2} n^2 \Rightarrow \frac{0.37}{a^2}, \tag{9}$$

(here the width of the well is expressed in nm), the observed frequency of jumps $I(U)$ should be expected for narrow wells and/or high excited states.

The blurred nature of the maxima observed in Fig.9 may be the result of imperfections in the structure described above, differences in the values of geometric and energy characteristics of the system, as well as the temperature factor. In the case of the presence of a magnetic field, there is a Zeeman effect, which in turn non-trivially changes the electronic states of the system, and therefore the dependence of $I(U)$.

CONCLUSION

As a result of the study, supramolecular clathrate structures were formed based on a 5-fold expanded GaSe matrix with an intercalated β-CD<FC> complex, featuring a hierarchical architecture of the subhost<host<guest>> type. Intercalation was performed under two technological conditions: under normal conditions and in the presence of a constant magnetic field. Based on the analysis of electrophysical measurements, the following conclusions were drawn:

1. Intercalation of the β-CD<FC> complex into the expanded GaSe structure leads to a substantial modification of the electrical conductivity, in particular an increase in resistance in the low-frequency range, an enhancement of the non-monotonic behaviour of the frequency dependence ReZ(ω), and the formation of additional trap levels associated with the emergence of quantum wells in the structure.
2. The synthesis conditions of the clathrate have a decisive effect on the electronic nature of the guest component: under normal conditions, the supramolecular complex β-CD<FC> exhibits acceptor-type conductivity, while intercalation in a constant magnetic

field induces donor properties. This is supported by the analysis of the frequency behavior of ReZ(ω) and the changes in the density of states near the Fermi level.

3. Analysis of the impurity energy spectrum parameters revealed that the introduction of the β-CD<FC> complex under normal conditions increases the density of states at the Fermi level and reduces the dispersion of trap levels near it. In contrast, the influence of the magnetic field leads to a nearly 10% decrease in the Fermi-level density of states N_F and a 10% increase in the hopping radius R.
4. The thermally stimulated discharge spectra indicate that the unexpanded GaSe crystal is characterised by a quasi-continuous distribution with heterocharge relaxation, whereas in the 5-fold expanded GaSe and the GaSe<β-CD<FC>> clathrates (regardless of synthesis conditions), the spectra exhibit a mini-zone structure with homocharge relaxation, suggesting the formation of trap centers localized within narrow bands in the bandgap.
5. The magnetoresistive and photoresistive effects of the GaSe<β-CD<FC>> clathrate are strongly dependent on the synthesis conditions. In particular, synthesis under a magnetic field results in a fivefold increase in magnetosensitivity and a twofold enhancement in photosensitivity, attributed to the ordering of the guest component and the specific configuration of impurity states.
6. The current-voltage characteristics (I-V curves) measured in a magnetic field for GaSe<β-CD<FC>> clathrates synthesized by both methods exhibit oscillatory behaviour, which may result from charge carrier tunnelling through quantum wells, Coulomb blockade, and the dynamic accumulation of nonequilibrium charge within the hierarchical subhost<host<guest>> architecture.

The results demonstrate the potential for targeted control over the electrical, magnetic, and optical properties of clathrate structures by tuning the intercalation conditions and guest component composition. This opens promising avenues for the development of sensor elements, photosensitive and magnetosensitive devices, and components for functional nanoelectronics.

COMPLIANCE WITH ETHICAL STANDARDS

The authors declare that the research was conducted in the absence of any potential conflict of interest.

AUTHOR CONTRIBUTIONS

Conceptualization, [F.I.]; methodology, [V.M.]; validation, [O.K., O.B.]; formal analysis, [N.P., O.B.]; investigation, [V.M]; resources, [R.S.]; data curation, [O.K.]; writing – original draft preparation, [F.I.]; writing – review and editing, [V.M.]; visualization, [F.I., V.M.]; project administration, [F.I.].

All authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] Ramamurthy, V., & Mondal, B. (2015). Supramolecular photochemistry concepts highlighted with select examples. *Journal of Photochemistry and Photobiology C: Photochemistry Reviews, 23*, 68–102. https://doi.org/10.1016/j.jphotochemrev.2015.04.002

[2] Amabilino, D. B., Smith, D. K., & Steed, J. W. (2017). Supramolecular materials. *Chemical Society Reviews, 46*, 2404–2420. https://doi.org/10.1039/C7CS00163K

[3] Hashim, P., Bergueiro, J., Meijer, E., & Aida, T. (2020). Supramolecular polymerization: A conceptual expansion for innovative materials. *Progress in Polymer Science, 105*, 101250. https://doi.org/10.1016/j.progpolymsci.2020.101250

[4] Hung, N. T., Nugraha, A. R. T., & Saito, R. (2017). Two-dimensional InSe as a potential thermoelectric material. *Applied Physics Letters, 111*, 092107. https://doi.org/10.1063/1.4999321
[5] Geim, A. K., & Grigorieva, I. V. (2013). Van der Waals heterostructures. *Nature, 499*, 419–425. https://doi.org/10.1038/nature12385
[6] Maksymych, V., Klapchuk, M., Borysiuk, A., Kulyk, Y., Stadnyk, V., Bordun, I., Kohut, Z., & Ivashchyshyn, F. (2023). Hierarchical heterostructure built on the basis of SiO_2 dielectric matrix and supramolecular complex β-cyclodextrin–ferrocene: Fabrication, physical properties and applications. *Materials Research Bulletin*, 164, 112220. https://doi.org/10.1016/j.materresbull.2023.112220
[7] Liu, G., Yang, J., & Xu, X. (2021). β-Cyclodextrin-calcium complex intercalated hydrotalcites as efficient catalyst for transesterification of glycerol. *Catalysts, 11*(11), 1307. https://doi.org/10.3390/catal11111307
[8] Yang, H., Chen, H., Chen, Z., Li, Y., Yao, L., Wang, G., Deng, Q., & Fu, P. (2023). Inductive effect of MXene membrane influenced by β-cyclodextrin intercalation. *Canadian Journal of Chemical Engineering, 101*(4), 1985–1992. https://doi.org/10.1002/cjce.24573
[9] Huang, T., Su, Z., Dai, Y., & Zhou, L. (2021). Enhancement of the heterogeneous adsorption and incorporation of uranium(VI) caused by the intercalation of β-cyclodextrin into the green rust. *Environmental Pollution, 290*, 118000. https://doi.org/10.1016/j.envpol.2021.118000
[10] Nair, A., Zhao, X., Glushenkov, A. M., et al. (2024). Metallocene intercalation in 2D materials: Recent advances and opportunities. *Journal of Materials Chemistry A, 12*, 1629–1650. https://doi.org/10.1039/D3TA07097B
[11] Gao, Y., Hu, G., Zhang, W., Ma, D., & Bao, X. (2011). π-π interaction intercalation of layered carbon materials with metallocene. *Dalton Transactions, 40*(17), 4521–4526. https://doi.org/10.1039/C0DT01392G
[12] Bal, B., Ganguli, S., & Bhattacharya, M. (1985). Intercalation of ferrocene in $CdPS_3$. *Physica B+C, 133*(1), 72–75. https://doi.org/10.1016/0378-4363(85)90026-9
[13] Harada, A., & Takahashi, S. (1984). Preparation and properties of cyclodextrin–ferrocene inclusion complexes. *Journal of the Chemical Society, Chemical Communications*, (10), 645–646. https://doi.org/10.1039/C39840000645
[14] Dupliak, I., Ivashchyshyn, F., Całus, D., Seredyuk, B., Chabecki, P., Maksymych, V., & Li, F. (2020). Influence of optical radiation and magnetic field on the properties of InSe<$NaNO_2$> clathrate. *Ukrainian Journal of Physics Optics, 21*(3), 115–122. https://doi.org/10.3116/16091833/21/3/115/2020
[15] Maksymych, V., Całus, D., Shvets, R., Chabecki, P., Pokladok, N., & Ivashchyshyn, F. (2022). GaSe<β-CD<I_2>> architecture supramolecular clathrate: Properties and application. *Journal of Nano- and Electronic Physics, 14*(1), 01002. https://doi.org/10.21272/JNEP.14(1).01002
[16] Barsoukov, E., & Macdonald, J. R. (2005). *Impedance spectroscopy: Theory, experiment and applications* (2nd ed.). Wiley.
[17] Pollak, M., & Geballe, T. H. (1961). Low-frequency conductivity due to hopping processes in silicon. *Physical Review, 122*(6), 1743–1753. https://doi.org/10.1103/PhysRev.122.1743
[18] Gyakushi, T., Asai, Y., Tsurumaki-Fukuchi, A., Arita, M., & Takahashi, Y. (2020). Periodic Coulomb blockade oscillations observed in single-layered Fe nanodot array. *Thin Solid Films*, *704*, 138012. https://doi.org/10.1016/j.tsf.2020.138012

МАГНІТО- ТА ФОТОКЕРОВАНІ СУПРАМОЛЕКУЛЯРНІ КЛАТРАТИ НА ОСНОВІ СЕЛЕНІДУ ГАЛІЮ

***Віталій Максимич*[1]*, *Олексій Корчев*[2], *Надія Покладок*[2], *Роман Швець*[2], *Олег Бордун*[1], *Федір Іващишин*[3]**

[1] Факультет електроніки та комп'ютерних технологій, Львівський національний університет імені Івана Франка, вул. Драгоманова, 50, 79005, м. Львів, Україна

[2] Інститут прикладної математики та фундаментальних наук, Національний університет «Львівська політехніка», вул. Степана Бандери, 12, 79013, м. Львів, Україна

[3]Електротехнічний факультет, Ченстоховський політехнічний університет, вул. Я.Г.Домбровського, 69, 42-201, м. Ченстохова, Польща

АНОТАЦІЯ

Вступ. Розвиток сучасної наноелектроніки та сенсорних технологій потребує створення нових функціональних матеріалів з керованими електричними, оптичними та магнітними властивостями. Одним із перспективних підходів є формування супрамолекулярних клатратів на основі шаруватих напівпровідників групи A^3B^6. У цій роботі представлено результати електрофізичного аналізу клатратних систем на основі GaSe з інтеркальованим комплексом β-циклодекстрин<фероцен> (β-CD<FC>), синтезованих за нормальних умов та в постійному магнітному полі.

Матеріали та методи. Як матриця-господар використовувався монокристал GaSe. Інтеркаляцію гостьового комплексу β-CD<FC> проводили у двох режимах: за нормальних умов і в постійному магнітному полі напруженістю 220 кА/м. Електрофізичні властивості досліджували методом імпедансної спектроскопії в діапазоні 10^{-3}–10^6 Гц та методом термостимульованого розряду.

Результати. Інтеркаляція β-CD<FC> у розширену структуру GaSe призводить до зростання опору в низькочастотному діапазоні та немонотонної поведінки ReZ(ω). Для зразків, синтезованих у магнітному полі, спостерігається зменшення ReZ(ω), що вказує на перехід гостьового компонента від акцепторного до донорного типу. Спектри термостимульованого розряду демонструють перехід від квазінеперервної до мінізонної структури рівнів з переважанням релаксації гомозаряду в клатратах. У зразках, синтезованих у магнітному полі, спостерігається п'ятикратне зростання магніторезистивного ефекту та збільшення фоточутливості вдвічі.

Висновки. Проведене дослідження продемонструвало, що інтеркаляція супрамолекулярного комплексу β-CD<FC> у розширену матрицю GaSe дозволяє цілеспрямовано модифікувати домішкову структуру і тим самим змінювати електропровідність та сенсорні властивості клатрату. Встановлено, що умови синтезу, зокрема дія постійного магнітного поля, суттєво впливають на електронну природу гостьового компонента, тип носіїв заряду, параметри пасткових рівнів і прояв квантових ефектів при перенесенні зарядів.

Ключові слова: Інтеркаляція, імпедансна спектроскопія, магніторезистивний ефект, фоторезистивний ефект, супрамолекулярний комплекс, ієрархічна архітектура.

Received / Одержано	Revised / Доопрацьовано	Accepted / Прийнято	Published / Опубліковано
09 May, 2025	30 May, 2025	05 June, 2025	26 June, 2025

Electronics and information technologies, 2025, 30, 177–186
http://publications.lnu.edu.ua/collections/index.php/electronics/index

UDC 535.323, 535.53, 537.226, 548

REFRACTOMETRY OF UNIAXIALLY COMPRESSED Rb_2SO_4 CRYSTALS

***Vasyl Stadnyk*[1]*, *Vasyl Vyshnevskyi*[1], *Igor Matviishyn*[2], *Vasyl Baliha*[1], *Roman Lys*[2]**
Ivan Franko National University of Lviv.
[1]Faculty of Physics, 19 Drahomanova Str., Lviv 79005, Ukraine.
[2]Faculty of Electronics and Computer Technologies, 107 Tarnavskoho Str., Lviv 79017, Ukraine.

Stadnyk, V. et. al (2025). Refractometry of uniaxially compressed Rb_2SO_4 crystals. *Electronics and Information Technologies*, 30. 177–186. https://doi.org/10.30970/eli.30.14

ABSTRACT

Introduction. The refractive index n of a crystal is its fundamental characteristic. In the presence of the experimentally obtained dispersion $n(\lambda)$ with high accuracy, it is possible to determine the fundamental characteristics of single crystals using analytical relations for the refractive index in the region of crystal transparency.

Materials and Methods. In this work, high-optical-quality rubidium sulfate (Rb_2SO_4) single crystals were synthesized, crystallographic orientation was determined using a polarization microscope and conoscopic figures, and the dispersion of the refractive indices $n_x(\lambda)$ was studied using the Obreimov immersion method at room temperature for different directions of uniaxial compression with stress magnitude of σ_m ~ 100 bar.

Results and Discussion. It was found that the refractive indices dispersion of this crystal $n_i(\lambda)$ is normal ($\partial n/\partial\lambda < 0$). As the fundamental absorption edge is approached, it increases sharply for the three polarization directions of the electromagnetic wave. The refractive indices satisfy the inequalities $n_x \geq n_z > n_y$ and $dn_x/d\lambda > dn_z/d\lambda > dn_y/d\lambda$.

The baric changes in the molar refraction R_i, electronic polarizability α_i (for the wavelength λ = 500 nm), and crystal-optical parameters (spectral positions of the effective centers of ultraviolet and infrared oscillators, as well as their effective strengths) of Rb_2SO_4 crystals at room temperature were calculated. It was found that uniaxial pressures of σ ~ 100 bar led to an increase in electronic polarizability of the crystal by an average of $(1 – 3)\times10^{-26}$ cm^3.

Conclusion. It was established that uniaxial pressures lead to an increase in contributions from infrared oscillators by 2–5 %, primarily due to the baric dependence of the optical isotropic point. They also shift the position of the effective strength of the ultraviolet oscillator λ_{0i} toward the short-wavelength region of the spectrum. The change in the effective oscillator strength in this baric range does not exceed 1 %. It was found that the positions of effective absorption bands are more sensitive to uniaxial pressures than the effective strengths of the corresponding oscillators.

Keywords: refractive indices, uniaxial compression, refraction, electronic polarizability, band gap width.

INTRODUCTION

The refractive index (RI) *n* of a crystal, which is determined by its electronic subsystem for a certain photon energy $\hbar\omega$ or wavelength λ, is its fundamental characteristic. The spectral dependence of the refractive index $n(\lambda)$ in a wide spectral region of the material's transparency contains significantly more information about its optical-electronic parameters.

The most accurate RI is determined for transparent materials in the visible and near ultraviolet (UV) and infrared (IR) regions of the spectrum with an error of the order of ±2×10^{-5}, but in several individual cases it can be ±5×10^{-6} [1-8]. Transparent materials in the specified regions of the spectrum include dielectrics and semiconductors with average bandgap widths (E_g ~ 4…6 eV). In the presence of the experimentally obtained dispersion $n(\lambda)$ with high accuracy, it is possible to determine the fundamental characteristics using analytical relations for the RI in the region of crystal transparency.

To describe the dependence of $n(\lambda)$, theoretical models of the atomic structure of the crystal are used, in which there is a dependence of the polarization of the medium on the frequency of the electromagnetic wave. If attenuation, nonlinearity, and spatial dispersion are neglected, then we can assume that the electrons behave similarly to the influence of the quasi-elastic force $\mathbf{F} = -q\mathbf{r}$ on their deviation from the equilibrium position:

$$r = \frac{eE}{m(\omega_0^2 - \omega^2)}, \tag{1}$$

where ω is the cyclic frequency of the electromagnetic wave $\mathbf{E} = \mathbf{E_0}\exp(-i\omega t)$, $\omega_0 = \sqrt{q/m}$ – the resonant frequency.

Each electron creates a dipole moment $\mathbf{p} = e\mathbf{r}$. The nuclei also contribute to the polarization of the crystal, but since their mass is large compared to the mass of the electron, their contribution in the first approximation can be neglected. The frequency dependence of RI is as follows:

$$\frac{n^2 - 1}{n^2 + 2} = \frac{4\pi}{3} N \frac{e^2}{m} \sum_k \frac{f_k}{\omega_k^2 - \omega^2}, \tag{2}$$

where N is the number of electrons for the corresponding resonant frequency ω_k, f_k is the oscillator strength, which determines the degree of electron participation in this oscillation.

If in relation (2) we consider only two oscillators in the near UV and IR regions of the spectrum, then we can obtain the well-known Selmeyer formula:

$$n_i^2 = 1 + \frac{B_1 \lambda_{01}^2 \lambda^2}{\lambda^2 - \lambda_{01}^2} + \frac{B_2 \lambda_{02}^2 \lambda^2}{\lambda_{02}^2 - \lambda^2}. \tag{3}$$

Here λ_{01} and λ_{02} are the spectral positions of the effective centers of the UV and IR oscillator bands, respectively, and B_1 and B_2 are the effective strengths of the corresponding oscillators.

The study of baric changes in the RI is due to the study of existing or possible phase transitions, as well as changes in the structure during temperature changes or compression of the crystal. Baric dependences of the RI have recently been actively studied, although the emphasis is on the influence of hydrostatic pressure [9 – 16].

It has been found that the RIs decrease on average by 4 % in the pressure range up to 15 kbar [5, 6] for several crystals that are actively used in multiplet quantum laser sources (InGaAsP, AlGaInAs, and GaInNAs). For sapphire crystals at pressures up to 190 kbar, the refractive index at a light wavelength of 514.5 nm decreases linearly [11]. For $KMgF_3$ crystals, the refractive index increases monotonically under pressure: dn/dp = 2.46×10^{-4} bar^{-1} for λ = 589.3 nm [12]. In some crystalline dielectrics, it has been found that the RIs mostly increase during uniaxial compression [13 – 16].

This behavior of RI is caused by the fact that hydrostatic pressure reduces the volume of the crystal and increases its density, which causes the growth of RI. The effect of uniaxial pressure on the refractive properties of crystals can be similar. At the same time, in anisotropic crystals, uniaxial compression along a certain direction will not always lead to unambiguous changes in the volume and RI in other crystallo-physical directions. Therefore, it is interesting to evaluate the effect of uniaxial compression on the RI of anisotropic crystals.

MATERIALS AND METHODS

We consider the effect of uniaxial compression σ_m on the behavior of the RI of a rubidium sulfate (RS) Rb_2SO_4 crystal. This is a typical representative of the $ABSO_4$ group crystals; at a temperature of T_c ~ 922 K, it undergoes a phase transition from the pseudohexagonal paraelectric phase *P*–3*m*1 to the orthorhombic ferro-elastic phase (symmetry space group $D_{2h}^{16} - Pmcn$) [17, 18]. The study of their crystal structure allowed us to clarify the lattice parameters: $a = 7.82079(10)$ Å, $b = 5.97778(7)$ Å, and $c = 10.44040(13)$ Å, and the band-energy structure` of the crystal has a direct-type band gap ($E_g = 4.89$ eV) and weak dispersion of energy levels. It was also established that for this crystal, the top of the valence band is formed by 2*p* states of oxygen atoms, and the bottom of the conduction band is formed by 4*s*-electrons of rubidium [17, 18].

Previously, an optical isotropic point (OIP) was found in this crystal at room temperature: the equality of n_z and n_y at a wavelength of λ = 495 nm and room temperature ($n_z = n_y = 1.51705$) [19-21]. The study of the dispersion dependences of birefringence $\Delta n_i(\lambda)$ showed that the crystal has a small normal dispersion $d(\Delta nx)/d\lambda \leq 0$, and uniaxial compressions σ_m do not change the character, but only the magnitude of the slope of the curves $\Delta n_i(\lambda)$ [22]. The authors also calculated the spectral dependences of the combined piezocoefficients $\pi_{im}^0(\lambda)$ and found that they are piezocoefficients π_{31}^0 and π_{21}^0 in the vicinity of the OIP at a wavelength of λ = 490 nm are equal.

The analysis of baric changes in the main RIs $n_i(\sigma_m)$ of the RS crystal was carried out based on previously obtained dependencies $\pi_{im}^0(\lambda)$ using the formula:

$$n_i(\lambda,\ T) = n_{io}(\lambda,\ T) - \frac{1}{2}\pi_{im}(\lambda,\ T)\sigma_m n_{io}^3(\lambda,\ T), \tag{4}$$

where $\boldsymbol{n}_{i0}$ is the refractive index of a mechanically free crystal,
σ_m is the magnitude of the mechanical load applied to the crystal,
π_{im} are the absolute piezo-optical constants,
index i denotes the direction of light,
index m denotes the direction of compression.

This research method allows us to study the behavior of all the main RIs under the action of pressures in a wide spectral region along different crystallographic directions (especially the behavior of $\boldsymbol{n}_i$ under the action of σ_i). The baric behavior of RIs $\boldsymbol{n}_i$ ($i \neq m$) can be directly investigated by the well-known direct interference-optical method [2, 3], but only for one wavelength, and it is experimentally more complicated.

RESULTS AND DISCUSSION

Fig. 1 shows the dispersion dependences of n_x at room temperature of mechanically free and uniaxially compressed RS crystals. In general, it was established that the dispersion of the RI $n_i(\lambda)$ of this crystal is normal ($\partial n/\partial\lambda < 0$), and increases rapidly with the approach to the edge of fundamental absorption, $n_i(\lambda)$ for three directions of

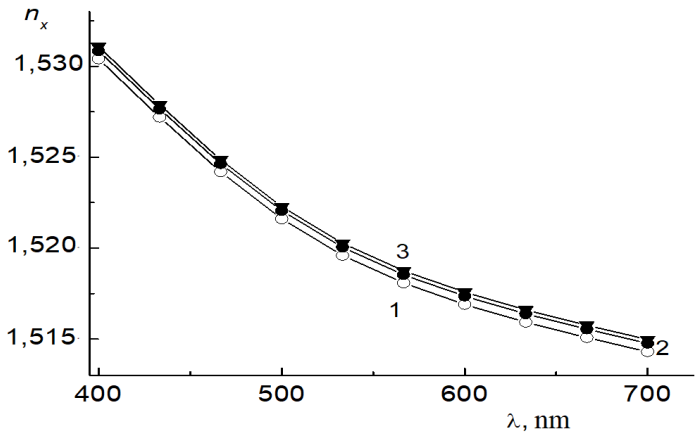


Fig. 1. Dispersion of refractive indices $n_x(\lambda)$ of Rb_2SO_4 crystals for different directions of uniaxial compression: 1 – σ_m = 0, 2 – σ_z = 100 bar, 3 – σ_y = 100 bar.

polarization of the electromagnetic wave. The following relations between the RI were found: $n_x \geq n_z > n_y$ and $dn_x/d\lambda > dn_z/d\lambda > dn_y/d\lambda$. It was established that the uniaxial pressure does not change the character of the curves $n_i(\lambda)$, only the value of the dispersion $dn_i/d\lambda$ changes. Table 1 shows the corresponding values for all crystallographic directions under the action of different pressures.

Let us consider the relationship $n_i(\lambda, \sigma)$ with the characteristics of the crystal, which follow from the dispersion formula of Selmeyer and the Lorentz-Lorentz formula, which describes refraction and electronic polarizability [1, 2]:

$$\alpha = \frac{3}{4\pi N}\frac{n^2-1}{n^2+2}, \tag{5}$$

Table 1. **Baric changes in the specific refraction R_i and electronic polarizability α_i for the light wavelength λ = 500 nm, as well as the crystal-optical parameters of Rb_2SO_4 crystals at a temperature of T = 293 K**

Parameters	σ_m = 0 bar	σ_x = 100 bar	σ_y = 100 bar	σ_z = 100 bar
$\alpha_x \times 10^{24}$, cm^3	8.85	8.87	8.88	8.88
$\alpha_y \times 10^{24}$, cm^3	8.83	8.85	8.85	8.86
$\alpha_z \times 10^{24}$, cm^3	8.84	8.86	8.87	8.86
R_x, cm^3	22.34	22.37	22.36	22.35
R_y, cm^3	22.30	22.31	22.32	22.33
R_z, cm^3	22.31	22.34	22.33	22.33
λ_{0x}, nm	94.19	94.25	94.23	94.23
λ_{0y}, nm	93.41	93.43	93.45	93.46
λ_{0z}, nm	94.18	94.19	94.21	94.22
$B_x \times 10^6$, nm^{-2}	141.15	141.13	141.12	141.12
$B_y \times 10^6$, nm^{-2}	143.83	143.81	143.82	143.80
$B_z \times 10^6$, nm^{-2}	141.16	141.13	141.14	141.12

$$R = \frac{4\pi}{3} N_A \alpha = \frac{\mu}{\rho} \frac{n^2 - 1}{n^2 + 2}. \quad (6)$$

Here N is the number of particles per unit volume; N_A is Avogadro's number; μ is the molar mass; ρ is the density of the crystal.

Using the obtained dependences $n_i(\lambda, \sigma)$, according to formulas (3), (5) and (6), the parameters listed in Table 1 were calculated. The calculated refractions of the RS crystals agree well with the sum of the individual refractions of the structural elements

$$R(Rb_2SO_4) = 2R(Rb) + R(SO_4) = 10 + 14.5 = 24.5\ cm^3/mol. \quad (7)$$

This suggests that Rb^+ cations make a significant contribution (~ 41 %) to the total refraction of the studied crystals. If we compare the crystal-optical parameters of these crystals with isomorphic ($(NH_4)_2SO_4$, $LiNH_4SO_4$, K_2SO_4 and $RbNH_4SO_4$), we can see that the cationic substitution $Li^+ \rightarrow Rb^+$ causes an increase in the RI on average by 3-5×10^{-3}, shifts the position of λ_{ox} by approximately 7 nm towards the short-wavelength, and the position of λ_{oz} – by 3 nm towards the long-wavelength parts of the spectrum. In general, it was established that the substitution $Li^+ \rightarrow K^+ \rightarrow NH_4 \rightarrow Rb^+$ leads to insignificant changes in λ_{01}, λ_{02}, B_1 and B_2 and a decrease in n_i by approximately 0.02 and 0.03.

It was found that the baric increase in the absolute value of the RI is due to a decrease in the effective power of the UV oscillator and a shift of the effective center of the UV absorption band to the long-wavelength region of the spectrum with a speed of $\partial\lambda_{0i}/\partial\sigma_m = 5 - 7 \times 10^{-3}$ nm/bar. If we differentiate (3), we can obtain a relation describing the baric changes in the RI:

$$\frac{dn_i}{d\sigma_m} = \frac{1}{2n_i}\left\{\frac{\lambda^2\lambda_{0i}}{\lambda^2 - \lambda_{0i}^2}\left(\frac{dB_1}{d\sigma_m}\lambda_{01} + \frac{2B_1\lambda^2}{\lambda^2 - \lambda_{01}^2}\frac{d\lambda_{01}}{d\sigma_m}\right) + \frac{\lambda^2\lambda_{02}}{\lambda^2{}_{02} - \lambda^2}\left(\frac{dB_2}{d\sigma_m}\lambda_{02} - \frac{2B_2\lambda^2}{\lambda^2{}_{02} - \lambda^2}\frac{d\lambda_{02}}{d\sigma_m}\right)\right\}. \quad (8)$$

It is considered that mechanical contact changes the forces of the oscillators, and this causes a shift of the corresponding effective centers of the absorption bands.

In this work, changes $d\lambda_{02}/d\sigma_m$ and $dB_2/d\sigma_m$ are neglected, since the effective center of the absorption band λ_{02} of the studied RS crystal is located in the far-IR region of the spectrum. Then, under the condition $dn_i/d\sigma_m > 0$, we can write:

$$\frac{d\lambda_{0i}}{d\sigma_m} > -\frac{\lambda_{0i}(\lambda^2 - \lambda_{0i}^2)}{2B_i\lambda^2}\frac{dB_i}{d\sigma_m}. \quad (9)$$

That is, by the nature of the baric changes in the RI, it is possible to establish the nature of the changes in the oscillators, which are decisive for the refractive properties of the crystal. In general, the anisotropy of electronic polarization corresponds to the anisotropy of the refractive indices ($\alpha_x > \alpha_z > \alpha_y$) (Table 1).

We see that pressures σ ~ 100 bar cause an increase in the parameter α_i on average by (1 – 3)×10^{-26} cm^3. This value coincides with the baric changes in the volume and linear dimensions of the sample along the direction of compression, which can be obtained based on Hooke's law:

$$\left(\frac{\Delta l}{l_0}\right)_i = s_{im}\sigma_m. \tag{10}$$

Here $(\Delta l/l_0)_i$ are the coefficients of relative elongation, and s_{im} are the coefficients of the elastic compliance tensor.

From formula (10), it is estimated that the coefficients $\Delta l/l_0$ ~ 10^{-4} for pressures σ_m ~ 100 bar and values of s_{im} ~ 10^{-6} bar^{-1}.

From the relation (6), we can obtain the following expression:

$$\frac{dR}{d\sigma} = -\left(\frac{\mu}{\rho^2}\frac{n^2-1}{n^2+2}\frac{d\rho}{d\sigma}\right)_n + \left[\frac{\mu}{\rho}\frac{6n}{(n^2+2)^2}\frac{dn}{d\sigma}\right]_N, \tag{11}$$

which allows us to find the contributions from the terms with derivatives $\partial\rho/\partial\sigma$ and $\partial n/\partial\sigma$ to the baric changes in the refraction R_i, provided that the values of n and N are constant during the corresponding experiment. If we use the known numerical data for the RS crystal: μ = 267 g/mol, ρ ~ 3613 kg/m^3, $\partial\rho/\partial\sigma \cong s_{im}$ = 0.5×10^{-10} Pa^{-1}, n_i = 1.52 and $\partial n/\partial\sigma \cong$ 5×10^{-6} bar^{-1}, we can obtain that the first term is approximately equal to 2×10^{-4} bar^{-1}, while the second is 4.8×10^{-4} bar^{-1}. This indicates that the geometric factor is ~ 28 % of the total baric increase in the refraction R.

In the case of neglecting the IR oscillator, the following contributions to the baric changes in the refractive indices can be obtained from relation (8):

$$\frac{\partial n}{\partial\sigma} \cong \frac{1}{2n}\cdot\frac{e^2}{\pi m c_0^2}\left\{\frac{\lambda^2\lambda_0^2}{\lambda^2-\lambda_0^2}\left[f\frac{\partial N}{\partial\sigma}+N\frac{\partial f}{\partial\sigma}\right]+\frac{2Nf\lambda_0\lambda^4}{(\lambda^2-\lambda_0^2)^2}\frac{\partial\lambda_0}{\partial\sigma}\right\} = \\ = 0.03\frac{\partial f}{\partial\sigma}+10^{-7}f+3\cdot10^{-6}f\frac{\partial\lambda_0}{\partial\sigma}. \tag{12}$$

Here it is taken into account that the coefficients B_1 and B_2 have the content of polarizabilities and are related to the oscillator strength by the following relations:

$$B_1 = \frac{N_{01}e_1^2f_1}{\pi c_0^2 m_1^*}; \qquad B_2 = \frac{N_{02}e_2^2f_2}{\pi c_0^2 m_2^*}. \tag{13}$$

Here N_{01} is the number of active particles per unit volume, $e_{1,2}$ and $m_{1,2}^*$ – the charges and effective masses of the oscillators, c_0 – the speed of light in vacuum. From the relation (12) we see that the main contribution to the baric changes of the RI is from the first term $\partial f/\partial\sigma$.

Baric changes of the RI $n_i(\sigma)$ make it possible to estimate the changes in the width of the forbidden band in accordance with the Moss relation [23]:

$$n^4E_g = const. \tag{14}$$

If we differentiate the last expression, we obtain:

$$\frac{\partial E_g}{\partial\sigma} = -\frac{4}{n}E_g\frac{\partial n}{\partial\sigma}. \tag{15}$$

If we use the obtained values of the baric changes n_i ($\partial n/\partial\sigma$ ~ 10^{-6} bar^{-1} (Fig. 1), n_i ~ 1.5 and E_g ~ 4.9 eV), we can estimate the corresponding changes in the width of the forbidden band $\partial E_g/\partial\sigma$ ~ 1.31×10^{-5} eV/bar.

The work also analyzes the baric changes λ_{0i}, B_{1i} and B_{2i} based on the baric dependences of the RI $n_i(\sigma)$ of the studied crystals. For this purpose, the formula (3) was differentiated, which allowed us to obtain the following relation:

$$\frac{dn_i(\lambda,\sigma)}{d\lambda} = -\frac{B_{1i}\lambda_{0i}^4\lambda}{n_i(\lambda^2-\lambda_{0i}^2)^2} - \frac{\lambda B_{2i}}{n_i}. \quad (16)$$

It was established that the contributions from UV oscillators to the dispersion changes $n_i(\lambda)$ of RS crystals decrease with increasing wavelength, and a significant anisotropy of the contributions from UV and IR oscillators was revealed. Thus, at room temperature, the contributions to the dispersion changes n_i from IR oscillators for λ = 500 nm are 68.1 % (X-direction) and 81.3 % (Z-direction). Under the action of uniaxial pressure, the contributions from UV oscillators decrease, and from IR – increase. For example, in the spectral region of 700 nm, in the case of increasing uniaxial load from 0 to 200 bar, the contributions from IR oscillators increase from 63.1 % to 66.4 % (Z-direction) (Table 2). At the same time, uniaxial pressures σ_m = 200 bar shift λ_{0i} to the short-wavelength region of the spectrum $\Delta\lambda_{0x}$ = 0.8…1.5 nm, $\Delta\lambda_{0z}$ = 1.3…1.9 nm.

The change in the effective oscillator strength in this pressure range does not exceed 1 %. Thus, the position of the effective absorption bands is more sensitive to the action of uniaxial pressures than the effective forces of the corresponding oscillators. The application of uniaxial pressures leads to a change in the contributions to the dispersion and temperature dependences $n_i(\lambda,T)$, and this change is not always unambiguous.

Uniaxial pressures cause an increase in contributions from IR oscillators by 2…5 %, which decreases with increasing temperature. This behavior of contributions from IR and UV under the action of different mechanical loads is due, first, to the baric dependence of the optical isotropic point. It was previously shown that for different directions of uniaxial compression, the optical isotropic point moves to a different spectral range, and its existence in any case is determined by the competition of IR and UV oscillators. In general, uniaxial pressures lead to a slight increase in contributions from IR oscillators, which decreases with increasing temperature. With increasing temperature, the increase in contributions from IR oscillators is insignificant – by 1–2 %. Uniaxial pressure σ_z leads to an increase in contributions from IR oscillators by 2–5 %, which decreases with increasing temperature.

The baric variation of the refractive indices of RS crystals is caused by structural changes in the crystal under the action of uniaxial compression. It is known that at room

Table 2. **Percentage ratio between contributions from ultraviolet and infrared oscillators to spectral and baric changes in refractive indices of Rb_2SO_4 crystals**

	σ = 0 bar						σ = 200 bar					
	300 nm		500 nm		700 nm		300 nm		500 nm		700 nm	
	UV	IR	UV	IR	UV	IR	UV	IR	UV	IR	UV	IR
X	88.8	11.2	41.4	58.6	21.3	78.7	84.3	15.7	39.4	60.1	20.2	79.8
Y	90.3	9.7	56.7	43.3	37.9	62.1	83.2	16.8	54.5	45.5	35.3	64.7
Z	88.7	11.3	55.4	44.6	36.9	63.1	81.7	18.3	55.8	44.2	33.6	66.4

temperature, the crystal structure is described by the space group symmetry *Pmcn*. In this phase, the tetrahedral SO_4 groups are somewhat distorted. The "average" spatial structure of RS at room temperature can be represented as follows: the basic orientational unit cell consists of SO_4^{2-} groups, each of which oscillates between three possible equilibrium configurations, defined by the position of the "apical" oxygen atom. The probability of a particular configuration depends solely on the oxygen atom's position, and the distortion of the two nonequivalent ions differs significantly. In general, the structure of the RS crystal in the paraelectric phase can be considered as a partially disordered state with respect to small rotations of SO_4 groups, mainly around the pseudo-hexagonal X-axis [18, 24].

CONCLUSION

In the work, Rb_2SO_4 crystals of good optical quality were synthesized, a crystallographic setup was set up, and the dispersion of the refractive indices $n_x(\lambda)$ at room temperature for different directions of uniaxial compression with a value of σ_m ~ 100 bar was studied using the Obreimov immersion method. It was found that the dispersion $n_i(\lambda)$ of this crystal is normal$\partial n/\partial\lambda < 0$ (), with an approach to the edge of the fundamental absorption $n_i(\lambda)$ it increases rapidly, and the following relations exist between the refractive indices: $n_x \geq n_z > n_y$ and $dn_x/d\lambda > dn_z/d\lambda > dn_y/d\lambda$.

The analysis of baric changes in the main refractive indices $n_i(\sigma_m)$ of a rubidium sulfate crystal was carried out and it was found that uniaxial pressure does not change the character of the curves $n_i(\lambda)$, but only the magnitude of the dispersion $\partial n_i/\partial\lambda$. It was found that the baric increase in the refractive index is mainly due to a change in the oscillator parameters (~ 72 %) due to a change in the band gap width E_g and a long-wave shift of the maximum of the ultraviolet absorption band and the density of effective oscillators (~ 28 %) of the crystals, as well as a decrease in the effective power of the ultraviolet oscillator and a shift of the effective center of the ultraviolet absorption band to the long-wave portion of the spectrum.

The baric changes in the specific refraction R_i, the electronic polarizability α_i (for the light wavelength λ = 500 nm), as well as the crystal-optical parameters (positions of the effective centers of the ultraviolet and infrared oscillator bands, as well as the effective forces of the corresponding oscillators) of Rb_2SO_4 crystals at room temperature were calculated. It was found that uniaxial pressures σ ~ 100 bar cause an increase in the electronic polarizability on average by (1–3)×10^{-26} cm^3, which coincides with the baric changes in the volume and linear dimensions of the sample along the compression direction. In general, the anisotropy of the electronic polarizability corresponds to the anisotropy of the refractive indices ($\alpha_x > \alpha_z > \alpha_y$).

Using the experimentally obtained values of the refractive indices, their baric changes, and the band gap width, the baric changes in the band gap width $\partial E_g/\partial\sigma$ were estimated to be ~ 1.31×10^{-5} eV/bar.

An analysis of the baric changes in the crystal-optical parameters was carried out and it was found that uniaxial pressures cause an increase in the contributions from infrared oscillators by 2…5 %, which is caused, first of all, by the baric dependence of the optical isotropic point, and also shift the position of the effective force of the ultraviolet oscillator λ_{0i} to the short-wavelength region of the spectrum. It was found that the position of the effective absorption bands is more sensitive to the action of uniaxial pressures than the effective forces of the corresponding oscillators.

COMPLIANCE WITH ETHICAL STANDARDS

The authors declare that they have no competing interests.

AUTHOR CONTRIBUTIONS

Conceptualization, [VS, VV, VB]; methodology, [VS, VV, IM]; investigation, [VS, VV, VB, RL]; writing – original draft preparation, [VS, RL]; writing – review and editing, [VS, VV, IM, VB, RL]; visualization, [VS, VV, VB].

All authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] Romaniuk M. O. Crystal optics. Lviv, Ivan Franko LNU, 2017, 456 p. (in Ukrainian).

[2] Romaniuk M. O., Krochuk A S., Pashuk I. P. Optics. Lviv, Ivan Franko LNU, 2012, 562 p. (in Ukrainian).

[3] Stadnyk V. Yo., Romaniuk M. O., Brezvin R. S. Electronic polarizability of ferroics: Monograph. Lviv, Ivan Franko LNU, 2014, 306 p. (in Ukrainian).

[4] Stadnyk V. Yo., Haba V. M. Refractometry of crystals with incommensurate phases: Monograph. Lviv, Liha-Pres, 352 p. (in Ukrainian).

[5] Shi Q., Dong L., Wang Y. (2023). Evaluating refractive index and birefringence of nonlinear optical crystals: Classical methods and new developments. *Chinese Journal of Structural Chemistry*, 42(1), 100017. DOI

[6] Wang S. Ren T., Yang P., Saito M., Brindley H. E. (2024). Improved Temperature-Dependent Ice Refractive Index Compilation in the Far-Infrared Spectrum. *Geophysical Research Letters*, 51(14), 1-10. DOI

[7] Ren T., Yang P., Brindley H. E., L'Ecuyer T. S., Maestri T. (2025). Temperature dependent optical properties of ice crysral in far-infrared regime. *Geophysical Research Letters*, 52(12), 1-11. DOI

[8] Herbin H., Deschutter L., Deguine A., Petiprez D. (2023). Complex refractive index of crystalline quartz particles from UV to thermal infrared. Aerosol Science and Technology, 57(3). 255-265. DOI

[9] Jin S. R., Sweeney S. J., Adams A. R., Higashi T., Riechert H., Thijs P. J. A. (2003). Wavelength dependence of the modal refractive index in 1.3 μm InGaAsP, AlGaInAs and GaInNAS lasers using high pressure. *Physica Status Solidi* (b), 235(2), 491-495. DOI

[10] Jones S. C., Vaughan B. A. M. and Gupta Y. M. (2001). Refractive indices of sapphire under elastic, uniaxial strain compression along the *a* axis. *Journal of Applied Physics*, 90(10), 4990–4996. DOI

[11] Jones S. C., Robinson M. C., Gupta Y. M. (2003). Ordinary refractive index of sapphire in uniaxial tension and compression along the *c* axis. *Journal of Applied Physics*, 93(2), 1023–1031. DOI

[12] Uchino K., Nomura S., Vedam K., Newnham R. E., Cross L. E. (1984). Pressure dependence of the refractive index and dielectric constant in a fluoroperovskite, $KMgF_3$. *Physical Review B*, 29, 6921–6925. DOI

[13] Fousek J. (1978). Reftactive indices and electro-optics at ferroelectrics and structural phase. *Ferroelectrics,* 20(1), 11–20. DOI

[14] Stadnyk V. Yo., Romanyuk M. O. (2005). The baric changes of the electron polarisability of $LiRbSO_4$, $LiKSO_4$ and $(NH_4)_2BeF_4$ crystals. *Ferroelectrics*, 317(1), 63–68. DOI

[15] Stadnyk V. Yo., Romanyuk M. O. (2005). The refractive properties of uniaxially stressed doped TGS crystals. *Ferroelectric*, 317(1), 95–99. DOI

[16] Stadnyk V. Yo., Romanyuk M. O., Chyzh O. Z.**,** Vachulovych V. F. (2007). The baric changes of the refractive properties of K_2SO_4 crystals. *Condensed Matter Physics*, 10(1), 45–60. DOI

[17] Muller O., Roy R. The Major ternary structural families. Berlin, Heidelberg and New York (Springer-Verlag), 1974, 487 p.

[18] Rudysh M. Ya., Pryshko I. A., Shchepanskyi P. A., Stadnyk V. Yo., Brezvin R. S., Kogut Z. O. (2022). Optical and electronic parameters of Rb_2SO_4 crystals. *Optik*, 269, 69875. DOI

[19] Gaba V.M. (2010). Temperature-and-Spectral Deformations of the Optical Indicatrix of Rubidium Sulphate Single Crystal. *Acta Physica Polonica A*, 117(1), 129-132. DOI

[20] Shchepanskyi P. A., Stadnyk V. Yo., Pryshko I. A., Brezvin R. S., Onufriv O. R., Kohut Z. (2024). Changes in thermal and refractive parameters of the sulfate group crystals in the region of the phase transition. *Journal of Physical Studies*, 28(1), 1701(1-7). DOI

[21] Pryshko I. A., Stadnyk V. Yo., Salapak V. M. (2023). On the low-temperature isotropic point in Rb_2SO_4 crystals. *Low temperature Physics*, 49(10), 1163–1169. DOI
[22] Stadnyk V., Matviishyn I., Ftomyn N., Vyshnevskyi V., Baliga V., Shtuka O. (2024). The photoelasticity of rhombic syngony crystals. *Electronics and information technologies*, 28, 157–167. DOI
[23] Narasimhampurty T. S. Photoelastic and Electro-Optic Properties of Crystals. New York, Springer, 1981, 544 p. DOI
[24] Rudysh M. Ya., Andriyevsky B. V., Shchepanskyi P. A. et al. (2025). Optical refractive properties and phonon spectra of Na_2SO_4 single crystal. *Optical Materials*, 166, 117228(10). DOI

РЕФРАКТОМЕТРІЯ ОДНОВІСНО ЗАТИСНУТИХ КРИСТАЛІВ Rb_2SO_4

***Василь Стадник*[1]*, Василь Вишневський*[1], Ігор Матвіїшин*[2], ***Василь Баліга*[1], Роман Лис*[1]
Львівський національний університет імені Івана Франка.
[1]Фізичний факультет, вул. Драгоманова 19, 79005 м. Львів, Україна.
[2]Факультет електроніки та комп'ютерних технологій, вул. Тарнавського 107, 79017 м. Львів, Україна.

АНОТАЦІЯ

Вступ. Показник заломлення (ПЗ) n кристалу є його фундаментальною характеристикою. За наявності експериментально отриманої з високою точністю дисперсії $n(\lambda)$ можна визначити фундаментальні характеристики з використанням аналітичних співвідношень для ПЗ в ділянці прозорості кристалу.

Матеріали та методи. У роботі синтезовано кристали Rb_2SO_4 високої оптичної якості, встановлено кристалографічну орієнтацію за допомогою поляризаційного мікроскопа та коноскопічних фігур, досліджено дисперсію показників заломлення $n_x(\lambda)$ методом занурення Обреїмова при кімнатній температурі для різних напрямків одновісного стиснення з величиною σ_m ~ 100 бар.

Результати. Виявлено, що дисперсія $n_i(\lambda)$ кристалу Rb_2SO_4 є нормальною ($\partial n/\partial\lambda < 0$), одновісний тиск не змінює характеру кривих $n_i(\lambda)$, а лише величину дисперсії $dn_i/d\lambda$. Встановлено, що баричне зростання показника заломлення, головно, зумовлено зміною параметрів осциляторів (~ 72 %) за рахунок зміни ширини забороненої зони E_g і довгохвильового зміщення максимуму смуги ультрафіолетового поглинання та густини ефективних осциляторів (~ 28 %) кристалів. Розраховано баричні зміни питомої рефракції R_i, електронної поляризовності α_i, а також кристалооптичних параметрів кристалів Rb_2SO_4 за кімнатної температури. Виявлено, що одновісні тиски σ ~ 100 бар спричиняють зростання електронної поляризовності в середньому на (1 – 3)×10^{-26} см^3, що співпадає з баричними змінами об'єму та лінійних розмірів зразка вздовж напряму стискання.

Висновки. Було встановлено, що одноосьові тиски призводять до збільшення внеску від інфрачервоних осциляторів на 2–5 %, насамперед через баричну залежність оптичної ізотропної точки. Вони також зміщують положення ефективної сили ультрафіолетового осцилятора λ_{0i} у бік короткохвильової області спектра. Зміна ефективної сили осцилятора в цьому баричному діапазоні не перевищує 1 %. Було виявлено, що положення ефективних смуг поглинання чутливіші до одноосьових тисків, ніж ефективні сили відповідних осциляторів.

Ключові слова: показники заломлення, одновісне стискання, рефракція, електронна поляризовність, ширина забороненої зони.

Received / Одержано	Revised / Доопрацьовано	Accepted / Прийнято	Published / Опубліковано
21 March, 2025	24 April, 2025	30 April, 2025	26 June, 2025

Збірник наукових праць

Електроніка та інформаційні технології

Electronics and information technologies

Випуск 30

2025

Підп. до друку 27.06.2025. Формат 70x100,16. Папір друк.
Друк на різогр. Гарнітура Times New Roman. Умовн. друк. арк. .
Тираж 100 прим. Зам. № .

Львівський національний університет імені Івана Франка.
79000 Львів, вул. Університетська, 1.

Свідоцтво про внесення суб'єкта видавничої справи до Державного реєстру видавців, виготівників і розповсюджувачів видавничої продукції. Серія ДК № 3059 від 13.12.2007 р.