Proceedings of the 6th Workshop for Young Scientists in Computer Science & Software Engineering (CS&SE@SW 2023). Virtual Event, Kryvyi Rih, Ukraine, February 2, 2024 / Edited by: Serhiy O. Semerikov, Andrii M. Striuk // CEUR Workshop Proceedings. – 2024. – Vol. 3662. – 259 p. – Access mode : https://ceur-ws.org/Vol-3662Наукову працю подано від установи "Інститут цифровізації освіти НАПН України"

Embracing Emerging Technologies: Insights from the 6th Workshop for Young Scientists in Computer Science & Software Engineering

Serhiy O. Semerikov[1,2,3,4,5], Andrii M. Striuk[4,1,5]

[1]Kryvyi Rih State Pedagogical University, 54 Universytetskyi Ave., Kryvyi Rih, 50086, Ukraine
[2]Institute for Digitalisation of Education of the NAES of Ukraine, 9 M. Berlynskoho Str., Kyiv, 04060, Ukraine
[3]Zhytomyr Polytechnic State University, 103 Chudnivsyka Str., Zhytomyr, 10005, Ukraine
[4]Kryvyi Rih National University, 11 Vitalii Matusevych Str., Kryvyi Rih, 50027, Ukraine
[5]Academy of Cognitive and Natural Sciences, 54 Gagarin Ave., Kryvyi Rih, 50086, Ukraine

Abstract

The 6th Workshop for Young Scientists in Computer Science & Software Engineering showcases cutting-edge research from emerging talents. This volume comprises diverse papers illuminating emerging technologies' profound impact across various domains. Several contributions underscore the pivotal role of telemetry, graph theory, and machine learning in optimising distributed systems, detecting anomalies, and streamlining processes. Others delve into acoustic surveillance techniques for UAV detection, genetic algorithms for university scheduling, and neural network-driven optimisation of chemical synthesis. The proceedings also highlight novel approaches to assessing software architecture reliability, implementing ERP systems, and designing information systems for viral infection data analysis. Thermal resistance calculation software, multimodal distribution data processing methods, and high-performance computing energy consumption modelling are also explored. Moreover, the importance of user experience research in cross-platform application development is emphasised, alongside the design of virtual physics laboratories and Python learning game applications. Notably, predatory conferences are addressed, proposing robust conference management platforms to uphold research integrity. Collectively, these papers exemplify young scientists' innovative spirit and determination to tackle real-world challenges and push the boundaries of their disciplines.

Keywords

emerging technologies, telemetry, graph theory, machine learning, acoustic surveillance, genetic algorithms, neural networks, software reliability, enterprise resource planning, user experience, virtual laboratories, Python learning games, predatory conferences

1. CS&SE@SW 2023: at a glance

Workshop for Young Scientists in Computer Science & Software Engineering (CS&SE@SW) is a peer-reviewed workshop focusing on research advances applications of information technolo-

CS&SE@SW 2023: 6th Workshop for Young Scientists in Computer Science & Software Engineering, February 2, 2024, Kryvyi Rih, Ukraine
semerikov@gmail.com (S. O. Semerikov); andrey.n.stryuk@gmail.com (A. M. Striuk)
https://kdpu.edu.ua/semerikov (S. O. Semerikov); https://scholar.google.com/citations?user=XzhtZZsAAAAJ (A. M. Striuk)
0000-0003-0789-0272 (S. O. Semerikov); 0000-0001-9240-1976 (A. M. Striuk)
© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

CEUR-WS.org/Vol-3662/paper00.pdf
CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

gies.

CS&SE@SW topics of interest since 2018 [1, 2, 3, 4, 5] are:

1. Software engineering
    - Software requirements [6, 7]
    - Software design [6, 8, 9, 7]
    - Software construction [10, 8, 9]
    - Software testing [6, 11]
    - Software maintenance [6]
    - Software engineering management [8]
    - Software development process [8, 9, 12, 7]
    - Software engineering models and methods [13, 10]
    - Software quality [14, 6, 11]
    - Software engineering professional practice [8]
2. Theoretical computer science
    - Data structures and algorithms [15, 16, 17, 9]
    - Theory of computation [15]
    - Information and coding theory [18, 19]
    - Formal methods [18]
3. Computer systems
    - Computer architecture and computer engineering [16, 17]
    - Computer performance analysis [16]
    - Databases [17]
4. Computer applications
    - Computer graphics and visualization [20, 12]
    - Human-computer interaction [21, 8, 17]
    - Scientific computing [20, 16, 17]
    - Artificial intelligence [22, 20, 13, 9, 19]

This volume represents the proceedings of the 6th Workshop for Young Scientists in Computer Science & Software Engineering (CS&SE@SW 2023), held in Kryvyi Rih, Ukraine, on February 2, 2024. It comprises 17 contributed papers that were carefully peer-reviewed and selected from 42 submissions. At least two program committee members reviewed each submission. The accepted papers present a state-of-the-art overview of successful cases and provide guidelines for future research.

2. CS&SE@SW 2023 Program Committee

- *Nadire Cavus*, Near East University [23, 24]
- *Stuart Charters*, Lincoln University [25, 26]
- *Attila Kertesz*, University of Szeged [27, 28]
- *Nagender Kumar Suryadevara*, University of Hyderabad [29, 30]
- *Orken Mamyrbaeyv*, Institute of Information and Computational Technologies [31, 32]
- *Bongkyo Moon*, QIR [33, 34]
- *Michael J. O'Grady*, University College Dublin [35, 36]
- *Grażyna Paliwoda-Pękosz*, Krakow University of Economics [37, 38]
- *Pedro Valderas*, Universitat Politècnica de València [39, 40]
- *Nataliia Veretennikova*, Lviv Polytechnic National University [41, 42]
- *Xianzhi Wang*, University of Technology Sydney [43, 44]
- *Alejandro Zunino*, ISISTAN - UNCPBA & CONICET [45, 46]

Additional reviewers:

- *Emrah Atilgan*, Eskişehir Osmangazi University [47, 48]
- *Olexander Barmak*, Khmelnytskyi National University [49, 50]
- *Kevin Matthe Caramancion*, University of Wisconsin–Stout [51, 52]
- *Pavlo Hryhoruk*, Khmelnytskyi National University [53, 54]
- *Oleksandr Kolgatin*, Simon Kuznets Kharkiv National University of Economics [55, 56]
- *Valerii Kontsedailo*, Inner Circle [57, 58]
- *Vyacheslav Kryzhanivskyy*, R&D Seco Tools AB [59, 60]
- *Andrey Kupin*, Kryvyi Rih National University [61, 62]
- *Mykhailo Medvediev*, ADA University [63, 64]
- *Vasyl Oleksiuk*, Ternopil Volodymyr Hnatiuk National Pedagogical University [65, 66]
- *Viacheslav Osadchyi*, Borys Grinchenko Kyiv University [67, 68]
- *James B. Procter*, University of Dundee [69, 70]
- *Serhiy Semerikov*, Kryvyi Rih State Pedagogical University [71, 72]
- *Etibar Seyidzade*, Baku Engineering University [73, 74]
- *Andrii Striuk*, Kryvyi Rih National University [75, 76]
- *Tetiana Vakaliuk*, Zhytomyr Polytechnic State University [77, 78]
- *Volodymyr Voytenko*, Athabasca University [79, 80]

3. CS&SE@SW 2023 organizers

The 6th edition of the CS&SE@SW was coordinated by the Academy of Cognitive and Natural Sciences (ACNS), a non-governmental organisation dedicated to nurturing the growth of researchers' expertise in the cognitive and natural sciences arena. ACNS's mission encompasses enhancing research, safeguarding rights and liberties, and catering to professional, scientific, social, and other interests.

ACNS is engaged in a spectrum of activities, including:

- Spearheading research initiatives within the cognitive and natural sciences domain and fostering collaborative ties among researchers.
- Orchestrating conferences, workshops, training sessions, internships, and other platforms for exchanging and disseminating knowledge in the realm of cognitive and natural sciences.
- Publishing conference proceedings, collections of scholarly works, and scientific journals (https://acnsci.org/journal):
  - Educational Dimension [81]
  - Educational Technology Quarterly [82]
  - CTE Workshop Proceedings [83]

Among ACNS's prominent publications is the Diamond Open Access *Journal of Edge Computing* (JEC), a peer-reviewed journal covering the science, theories, and practice of IoT, distributed systems, and edge computing [84]. JEC considers scientific research on using and applying edge computing in various fields: education, science, medicine, architecture, etc. [85]. Notably, JEC covers a broad range of topics aligned with CS&SE@SW topics of interest:

- Artificial intelligence [86, 87]
- Computer networks [88]
- Computer performance analysis [89]
- Concurrent, parallel and distributed systems [90, 91, 92]
- Formal methods [93]
- Human-computer interaction [94, 95, 89, 84]
- Mathematical foundations [96]
- Scientific computing [97, 98, 99, 100, 89]

4. CS&SE@SW 2023 keynote

This year, one keynote speaker was selected by the CS&SE@SW 2023 program committee: Dmytro Nechai (Chief architect at PLATMA, CTO at SalesJinn, mentor and lector at National Technical University of Ukraine and "Igor Sikorsky Kyiv Polytechnic Institute") "The future is already here. What is low-code and what to serve it with?" (figure 1).

5. CS&SE@SW 2023 articles overview

5.1. Software engineering

The article "An approach to assessing the reliability of software systems based on a graph model of method dependence" by Krutko et al. [14] proposes a method for evaluating the reliability of software systems. The authors highlight the importance of software quality, particularly reliability, in today's rapidly evolving software development landscape. They observe that existing reliability assessment methods often rely on hardware models, which may not fully capture the intricacies of software systems.

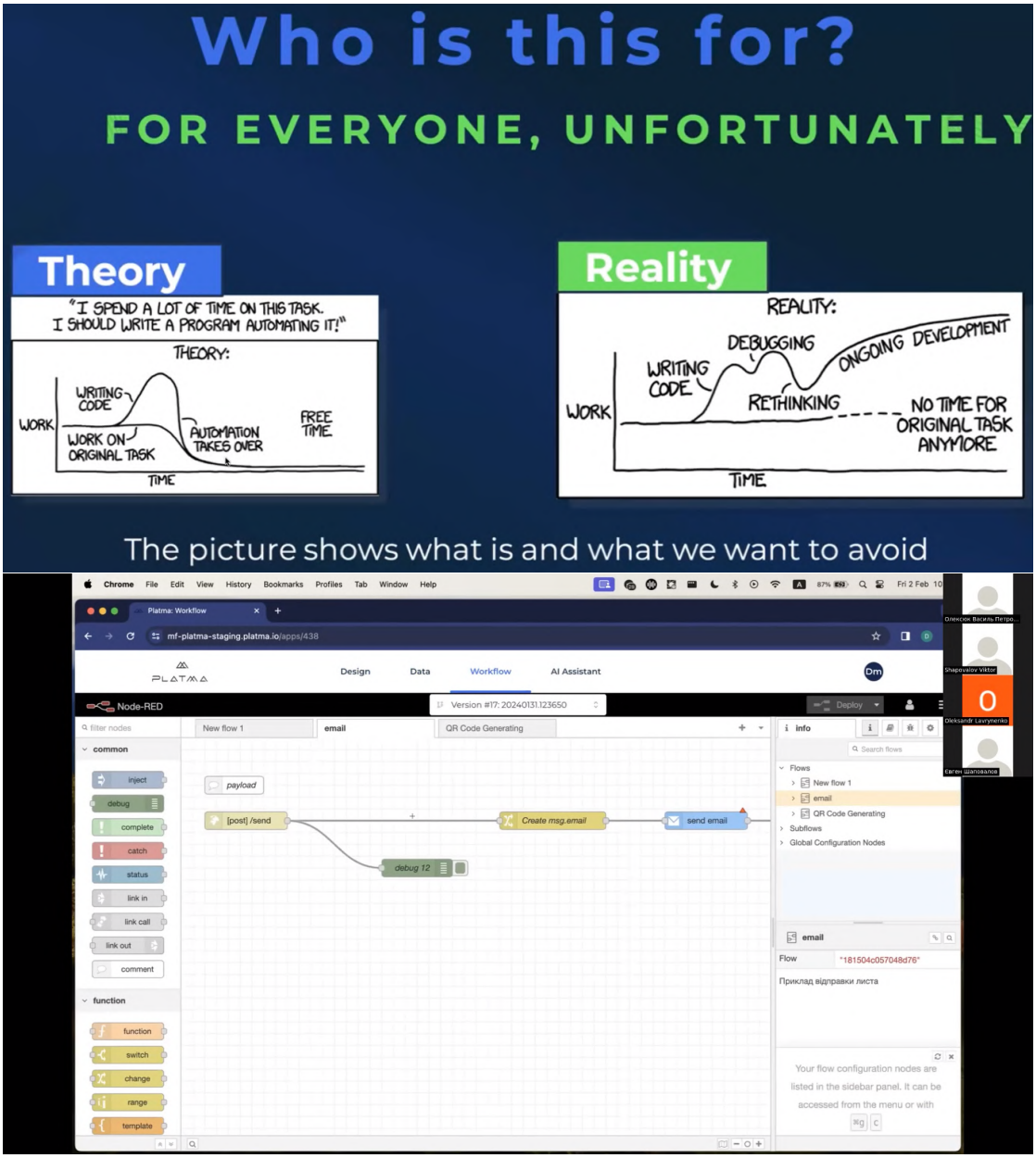


Figure 1: CS&SE@SW 2023 keynote.

The proposed approach introduces a graph model of method dependence, wherein software systems are broken down into smaller structural elements called methods. These methods are then analysed to construct a graph model representing their interdependencies. Stochastic reliability indicators are assigned to each method based on the probability of failure-free operation. These indicators are calculated by analysing method invocations and failures during program execution.

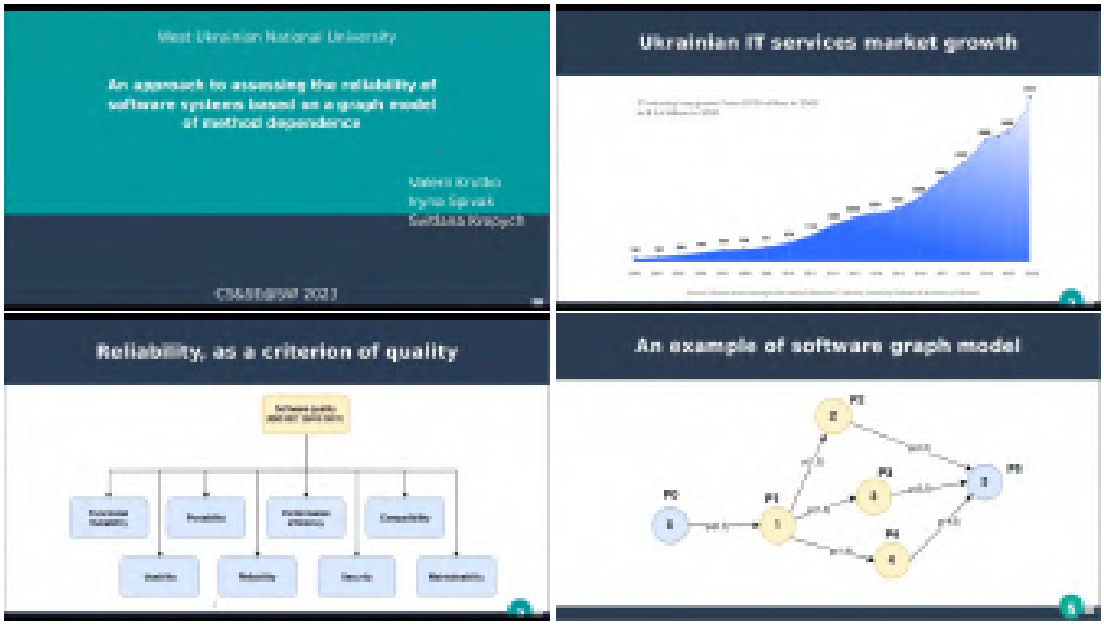


Figure 2: Presentation of paper [14].

The article describes the proposed method, including the steps in constructing the graph model and calculating reliability indicators. It also presents examples demonstrating the application of the approach to simple and complex software systems.

The article "Methodology of implementation of modern information systems at commercial enterprises" [6] provides a comprehensive overview of implementing ERP (Enterprise Resource Planning) systems based on the AIM (Application Implementation Method) methodology, with a focus on Ukrainian realities. Authored by Yurii O. Chernukha, Oksana V. Klochko, and Tetiana P. Zuziak from Vinnytsia Mykhailo Kotsiubynskyi State Pedagogical University, Ukraine, the article delves into the various stages of implementing ERP systems, including preparation and planning, selecting an ERP system, design, development and testing, training and support, analysis and optimisation, and support and updates.

The authors meticulously detail each phase, providing insights into the tasks, challenges, and considerations associated with implementing ERP systems. They emphasise the importance of careful planning, stakeholder collaboration, and continuous monitoring throughout the implementation process. Furthermore, they highlight the significance of selecting the right ERP system and project management strategies to ensure successful outcomes.

A notable aspect of the article is its discussion on the methodology for implementing ERP systems. It mainly focuses on the Oracle AIM methodology, which divides the project into six phases and encompasses various processes within each phase. The authors provide an in-depth analysis of the documents associated with each process, offering readers a comprehensive understanding of the documentation required for successful ERP implementation.

Moreover, the article addresses the challenges specific to Ukrainian enterprises, such as historical processes, diverse applications, and limited documentation, and provides practical

recommendations for overcoming these challenges. It emphasises the importance of organisational restructuring, business process optimisation, and the active involvement of company management in the implementation process.

Additionally, the article discusses the role of project management tools and communication platforms in facilitating collaboration and coordination among project teams. It highlights the significance of Microsoft Project, Jira, Confluence, and other tools in streamlining project activities and ensuring effective communication among team members.

The article "Information System Module for Analysis of Viral Infections Data Based on Machine Learning" [13] presents a comprehensive exploration of the development and implementation of an information system module designed to analyse viral infection data. Authored by Nickolay Rudnichenko, Vladimir Vychuzhanin, Tetiana Otradskya, and Igor Petrov, the article delves into the significance of automating data analysis processes, particularly in the context of viral diseases, utilising intelligent technologies and machine learning methods.

The article begins by addressing the relevance of data analysis automation in various fields, emphasising the importance of modern tools and approaches in efficiently handling large volumes of data. With a focus on viral diseases, especially in the post-COVID-19 era, the authors highlight the ongoing need for analysing disease patterns, forecasting, and automating symptom detection to prevent further spread.

Key components and technologies used in developing the information system module are described, including using the UML language for system design modelling, client-server architecture, and relational database implementation. The process of creating, training, and testing machine learning models is detailed, along with assessing input features' significance and error matrix evaluation.

The article provides insights into the project structure, outlining the system's functionalities such as authentication, dataset importation, data visualisation, and model parameter modification. It also presents a sequence diagram illustrating the system's operation and a component diagram highlighting its main modules.

Results from implementing five machine learning models – Gaussian Naive Bayes, Decision Tree, Random Forest, Support Vector Machine, and Neural Network – are discussed, along with the performance metrics and analysis of each model's outputs. The authors demonstrate the effectiveness of these models in analysing COVID-19 symptom data, identifying key symptoms indicative of the virus, and assessing model accuracy and speed.

In summary, the article provides a thorough overview of developing and implementing an information system module for analysing viral infection data using machine learning techniques. It underscores the importance of automated data analysis in addressing public health challenges, with implications for improving disease prevention and control strategies. The findings contribute to advancing research in the field of data-driven healthcare and highlight avenues for future exploration, including developing more efficient models and expanding datasets for comprehensive analysis.

The article "Designing a cross-platform user-friendly transport company application" [8] delves into the crucial aspects of developing an application for a transportation company with a focus on cross-platform compatibility user experience (UX) and user interface (UI) design. Authored by Maksym Y. Salohub, Olena H. Rybalchenko, and Svitlana V. Bilashenko from Kryvyi Rih National University, the paper presents a comprehensive approach to creating a scalable

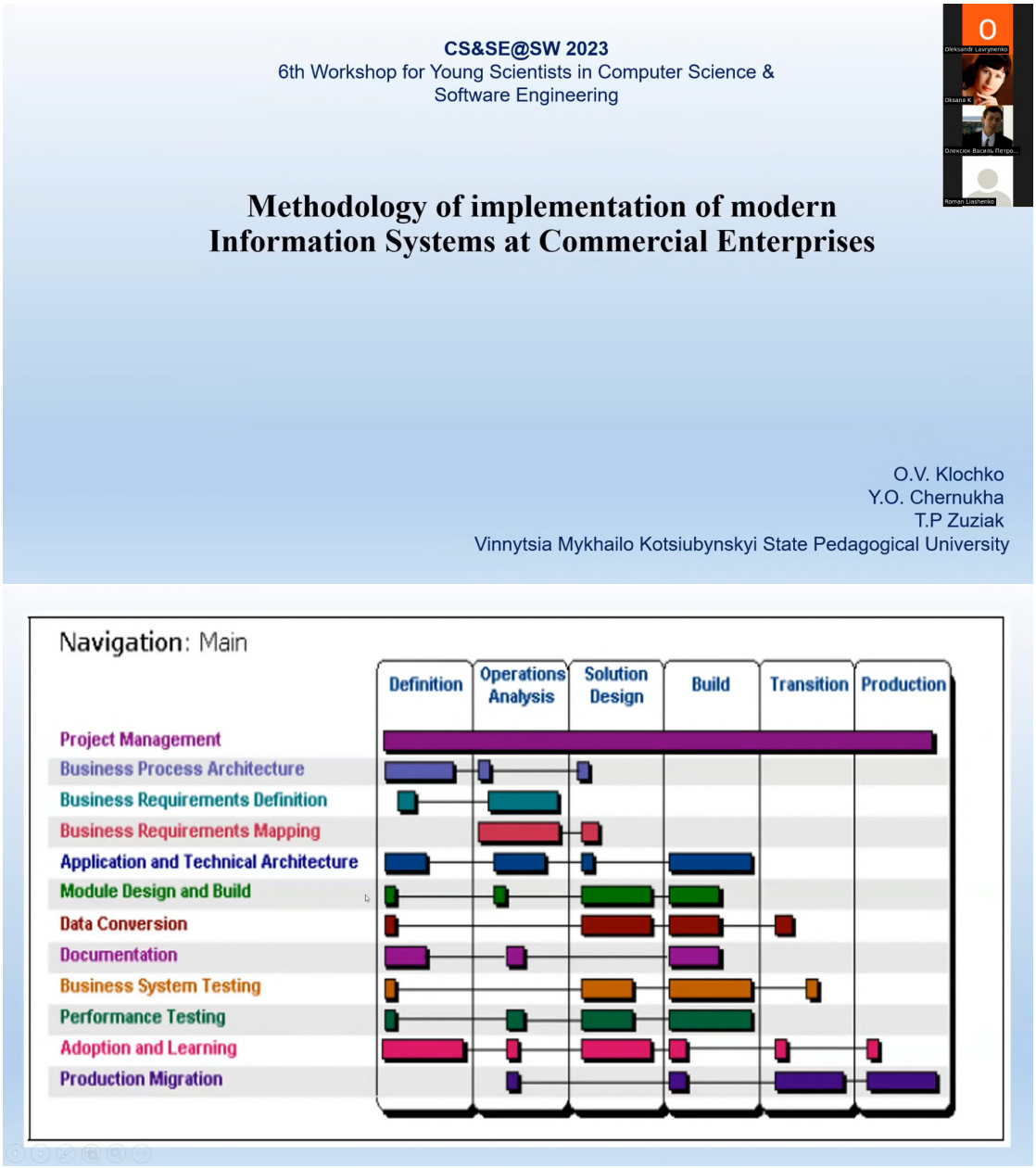


Figure 3: Presentation of paper [6].

and user-centric application.

The methodology includes UX research, competitor analysis, and target audience surveys to identify user preferences and behaviours. Through a thorough analysis of analogous applications like Bolt, Grab, and DiDi Rider, the authors provide insights into the strengths and weaknesses of existing platforms. Additionally, the survey results highlight the importance of features such

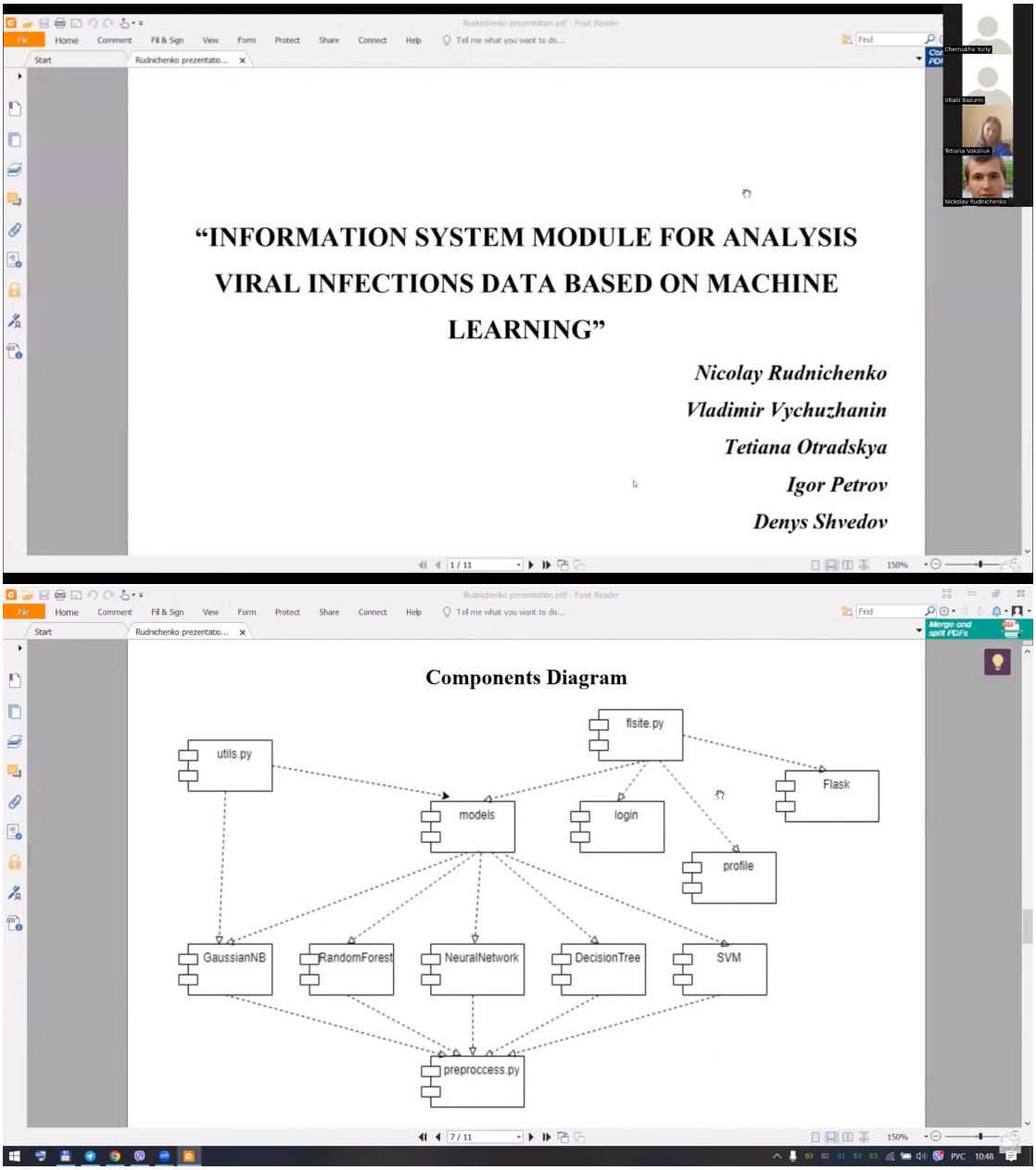


Figure 4: Presentation of paper [13].

as panic buttons, driver selection options, and trip archives for users.

The article discusses various approaches to cross-platform development, emphasising the advantages of using technologies like React Native to streamline the development process and ensure compatibility across different platforms. The authors also address challenges in UI design and propose solutions to create an intuitive and visually appealing interface.

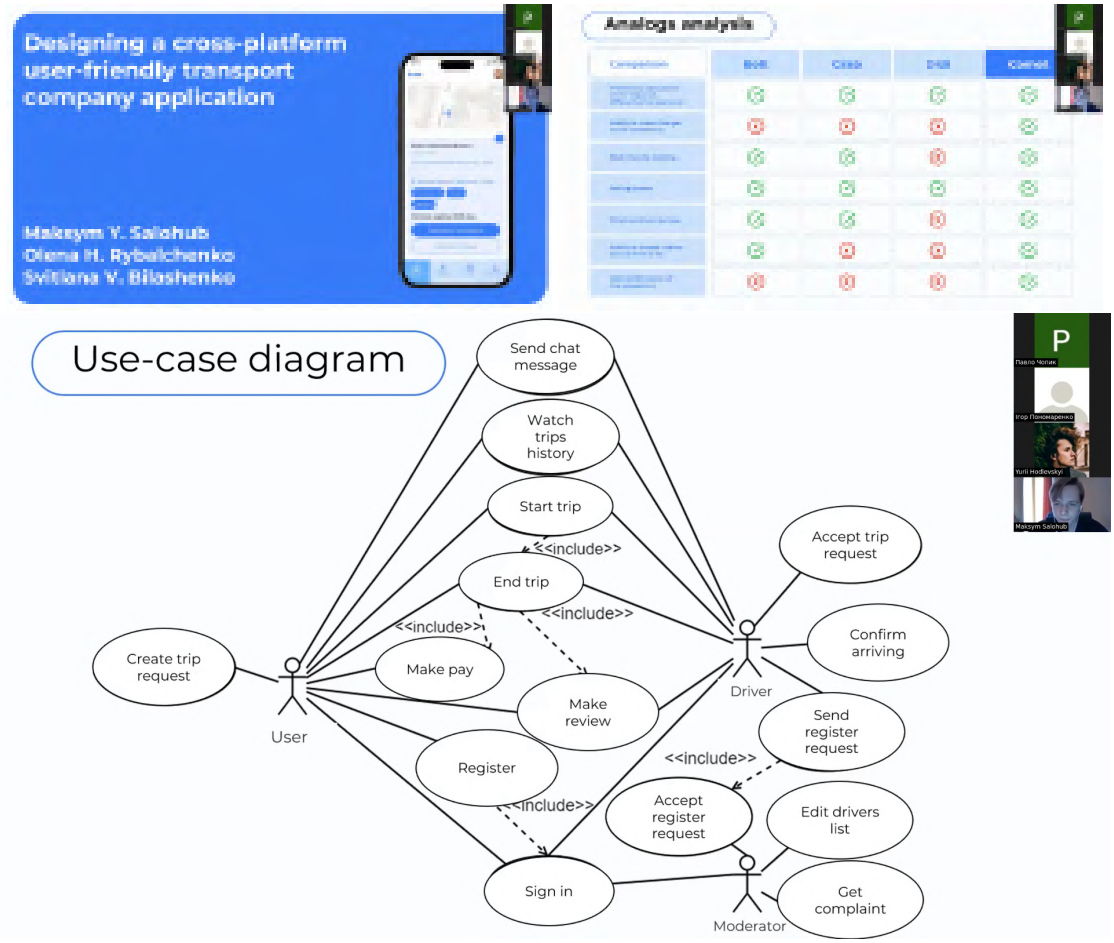


Figure 5: Presentation of paper [8].

Furthermore, the paper outlines the system development process, including using technologies like Express.js for backend development and MongoDB for database management. The integration of Expo CLI facilitates testing and deployment, while the utilisation of Feather Icons enhances the semantic interaction within the application.

The article “Research of the route planning algorithms on the example of a drone delivery system software development” [9] provides an in-depth analysis of various route planning algorithms for drone delivery systems. Authored by Yevhen L. Turchyk, Milana V. Puzino, Olena H. Rybalchenko, and Svitlana V. Bilashenko, the paper delves into the existing drone delivery systems worldwide, examines different route-building algorithms, and discusses the advantages and disadvantages of each approach.

The paper begins with an introduction highlighting the significance of efficient logistics, particularly in urban settings. It introduces the concept of drone delivery as a potential solution to overcome challenges in last-mile delivery. It sets the stage for the research by emphasising the need for quick and convenient operation in drone delivery systems.

The subsequent sections thoroughly review existing drone delivery systems, such as Amazon Prime Air, Starship Technologies, and Zipline, providing insights into their operations, advantages, and limitations. Recent research on drone delivery systems is also analysed, covering topics like multi-physics modelling, cloud-based drone management, and optimisation algorithms for route planning.

A comprehensive review of common approaches and algorithms for drone delivery route planning follows, including the Traveling Salesman Problem algorithm, Dijkstra's algorithm, A* algorithm, and reinforcement learning. Each algorithm is evaluated based on execution speed, scalability, and implementation simplicity. The authors argue that reinforcement learning emerges as the most optimal solution due to its ability to adapt to dynamic environments and optimise delivery routes efficiently.

The paper concludes with a discussion on system development, outlining the general architecture of a drone delivery system, hardware simulation using ArduPilot SITL, and implementing a route-building subprogram using Q-Learning. The provided code snippets offer insights into how reinforcement learning techniques can be applied to optimise delivery routes.

The article "Implementing E2E tests with Cypress and Page Object Model: evolution of approaches" [11] presents a comprehensive exploration of various methodologies for constructing Cypress tests using the Page Object Model (POM). Authored by Inessa V. Krasnokutska and Oleksandr S. Krasnokutskyi from Yuriy Fedkovych Chernivtsi National University, the article delves into different strategies for organising tests with Cypress while utilising the POM design pattern.

The authors begin by introducing the problem of automating tests for a website, using the example of the saucedemo.com website. They emphasise the importance of covering positive and negative test cases, such as successful logins and unsuccessful login attempts resulting in error messages.

The article outlines nine distinct approaches to implementing the Page Object Model with Cypress. These approaches range from tests without POM to utilising POM with various techniques, such as selectors for elements, getters for error messages, and assessor properties. Each approach is discussed in detail, highlighting its advantages, disadvantages, and evolution from simpler to more refined implementations.

The article provides code snippets and examples to illustrate each approach, making it accessible for readers to understand and implement in their projects. The authors also provide insights into the rationale behind each approach, discussing factors such as code maintainability, readability, and adherence to best practices.

The article "Design and development of a game application for learning Python" by Oleksiuk et al. [7] explores the creation of a Python learning game application and presents the outcomes of meeting its objectives. The authors analyse various game-based learning experiences, establish application requirements, select Unity3D as the game engine, and describe their experience in developing the PythonLeaner game.

The article begins by discussing the significance of game-based learning in teaching programming. It highlights its benefits, such as increased engagement, active participation, hands-on learning, and simulation of real-life scenarios. It then outlines the research objectives, including analysing experiences, describing application requirements, selecting development tools, and analysing key development points.

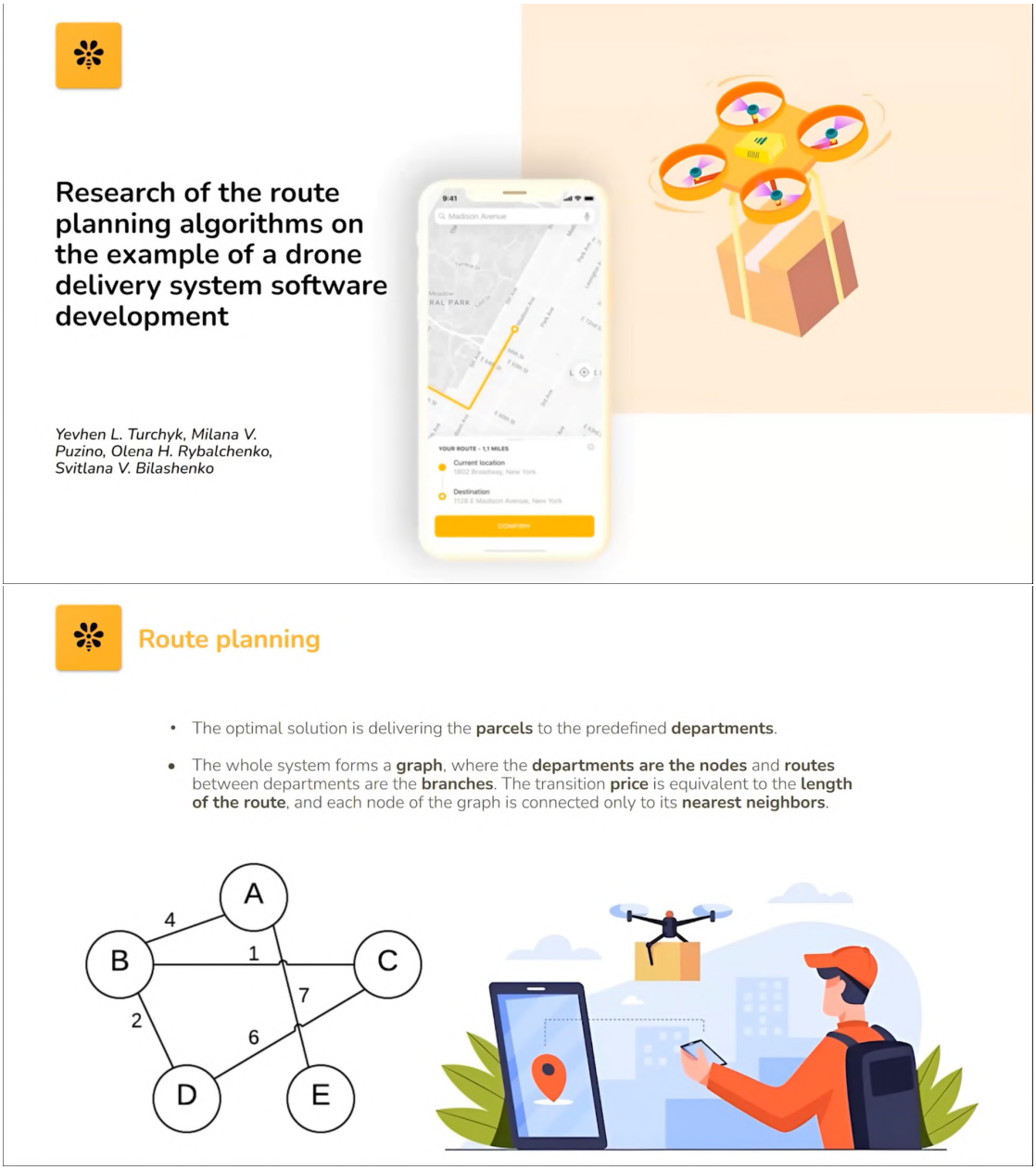


Figure 6: Presentation of paper [9].

The model of the game application for learning Python is described, emphasising the incorporation of educational objectives, game mechanics, hands-on learning, individualised progression, and reporting of learning outcomes. The game model includes modes such as New Game, Continue, Shop, and Exit, emphasising individualised progression through levels.

A comparison of game engines Unity3D, Unreal Engine, and CryEngine is provided, high-

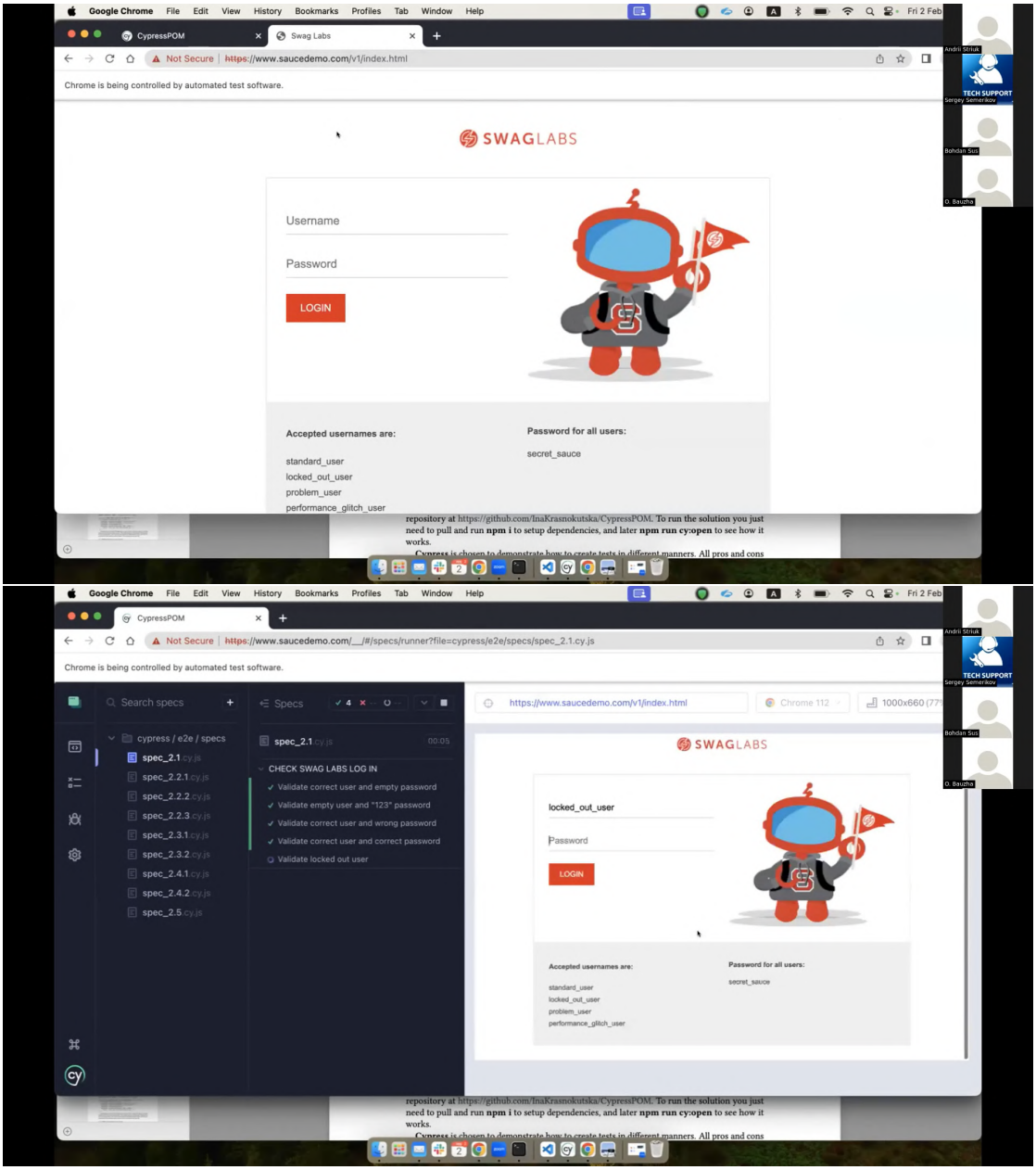


Figure 7: Presentation of paper [11].

lighting Unity3D as the chosen platform for its ease of learning, compatibility, multi-platform support, and active community. The article then analyses key development points, including scene design, script creation, Firebase integration for data storage, and implementation of game features such as animations, user input delay, and task types.

The conclusion summarises the achieved objectives, emphasising the analysis of experiences,

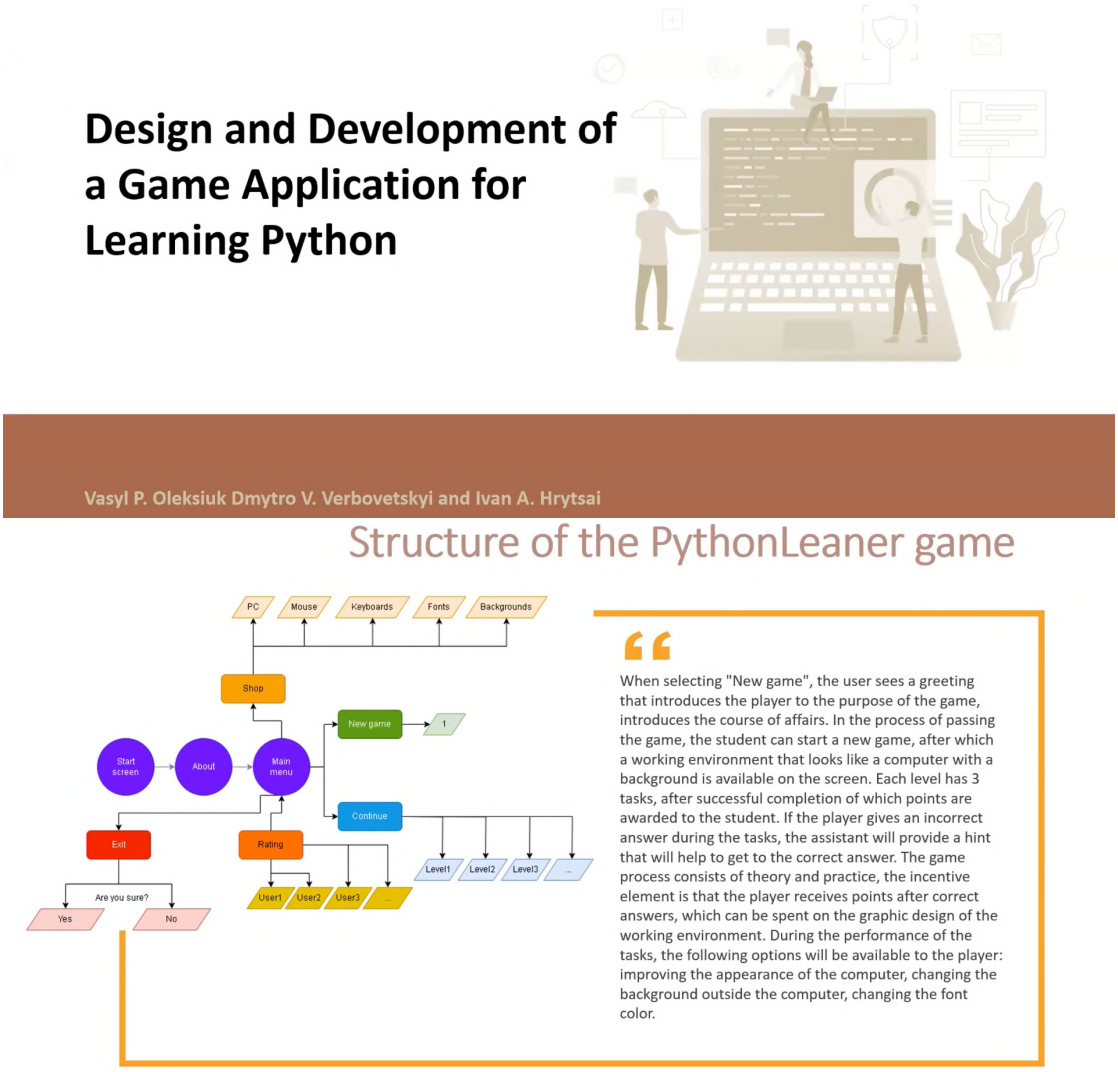


Figure 8: Presentation of paper [7].

the establishment of application requirements, the selection of Unity3D as the game engine, and the description of crucial development points. It also outlines prospects for research, including multiplayer integration, code interpretation, artificial intelligence, and mobile application development.

5.2. Theoretical computer science

The article “Application of Daubechies wavelet analysis in problems of acoustic detection of UAVs” [18] provides an in-depth exploration of the utilisation of Daubechies wavelet analysis for acoustic signal processing in the context of detecting unmanned aerial vehicles (UAVs). Authored by Oleksandr Yu. Lavrynenko et al. from the National Aviation University in Ukraine, the study

addresses the significance of acoustic surveillance in UAV detection. It proposes Daubechies wavelet analysis as a promising method for identifying characteristic features of UAVs' acoustic radiation. The article offers a thorough exploration of Daubechies wavelet analysis in the context of acoustic UAV detection, providing theoretical foundations and practical insights into the application of this method. It bridges the gap between theoretical wavelet analysis and its implementation in real-world problems, making it a valuable resource for researchers and practitioners in signal processing and UAV detection.

The article "Data processing method for multimodal distribution parameters estimation" by Solomentsev et al. [15] describes the synthesis and analysis of a method for processing data to estimate the parameters of multimodal distributions. The proposed approach combines the method of moments and the method of quantiles. The method allows for estimating the parameters of the probability density function even without prior information about the distribution type, which is essential in practical applications, especially in telecommunications and radio engineering.

The key steps of the method include dividing the sample population into subsets corresponding to positive and negative regions, selecting appropriate thresholds based on the distribution characteristics, and employing a combination of moment-based and quantile-based estimation techniques to estimate the parameters of interest. The approach is illustrated with a specific example of the trimodal probability density function, which includes chaotic impulse noise of positive and negative polarity.

The proposed method offers a practical solution for estimating distribution parameters in scenarios where the distribution type is unknown or complex. Future research could explore further refinements and extensions of the method and its application in various real-world data processing tasks.

The article "Application of artificial intelligence in digital marketing" [19] provides a comprehensive analysis of how artificial intelligence (AI) can be utilised to optimise digital marketing strategies for companies. Authored by Ihor V. Ponomarenko, Volodymyr M. Pavlenko, Oksana B. Morhulets, Dmytro V. Ponomarenko, and Nataliia M. Ukhnal, the paper explores various aspects of AI integration into digital marketing tools, emphasising its role in enhancing user engagement, personalisation, content generation, customer support, sentiment analysis, and more.

The authors begin by highlighting the significance of digitisation processes in reshaping consumer behaviour and increasing dependence on innovative technologies. They argue that AI catalyses qualitative transformations in digital marketing, enabling companies to leverage vast amounts of data generated online for strategic decision-making. Through a methodological approach grounded in scientific analysis, the paper outlines the primary sources of information utilised in AI applications for digital marketing, including data from company websites, social media, public sources, and web scraping.

Furthermore, the article delves into the models and methods employed in AI-driven digital marketing, emphasising the importance of data analysis, content personalisation, and customer interaction channels. It discusses the role of machine learning algorithms in processing big data, segmenting target audiences, generating personalised content, and enhancing customer support services. The authors also highlight the significance of sentiment analysis in gauging user attitudes and adjusting marketing strategies accordingly.

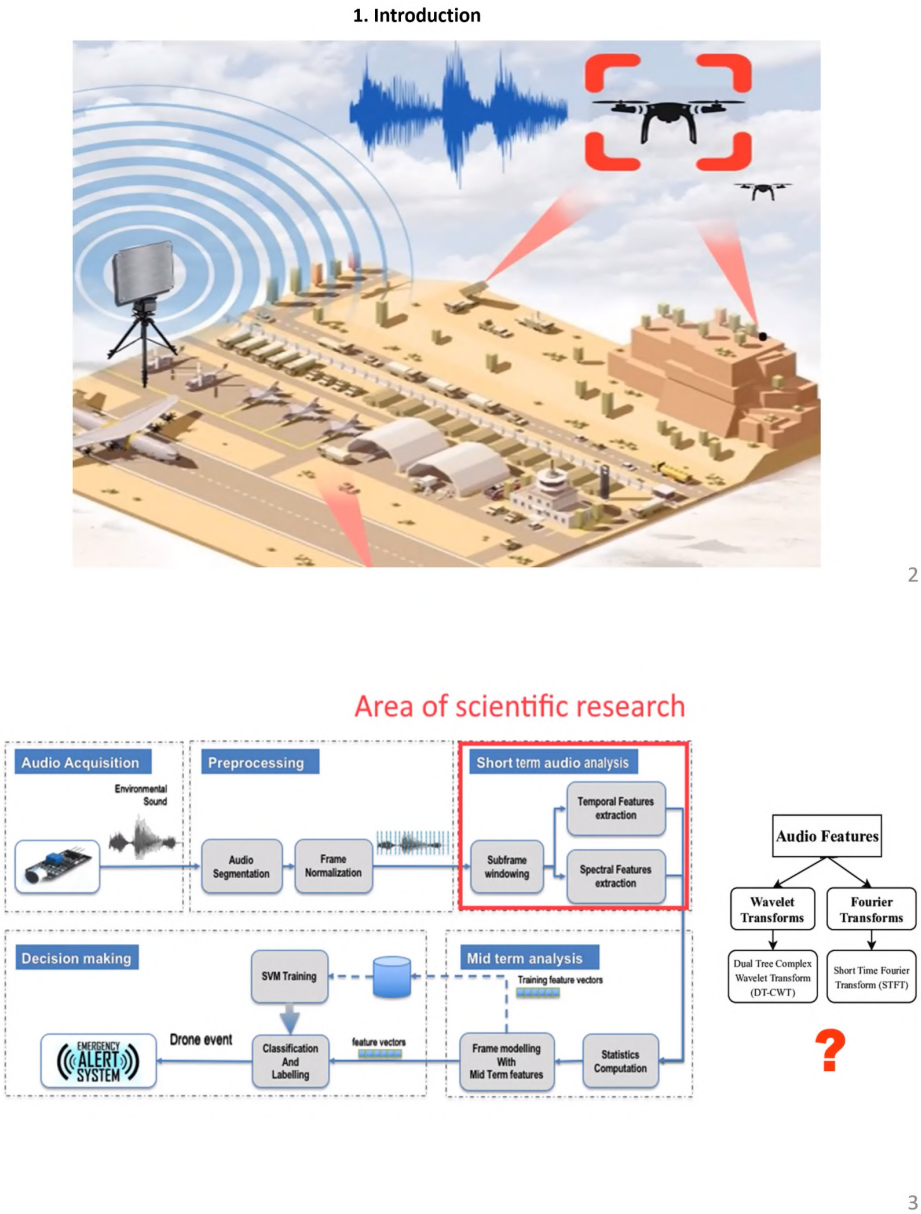


Figure 9: Presentation of paper [18].

In addition to providing insights into current practices, the article identifies future research directions in AI-driven digital marketing. It emphasises the need for ongoing development of machine learning algorithms, specialised programming languages, and innovative methodological approaches to optimise marketing strategies further and enhance user experiences.

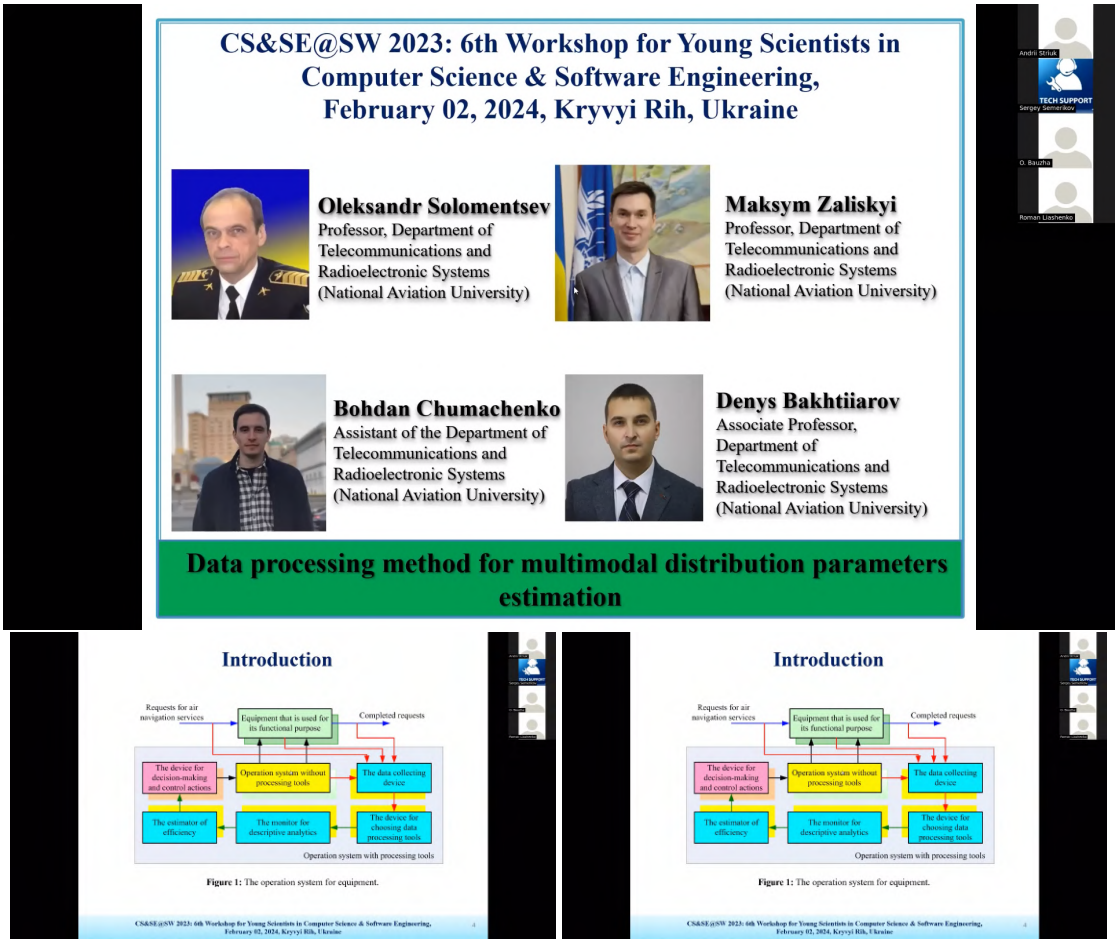


Figure 10: Presentation of paper [15].

5.3. Computer systems

The article "Modern methods of energy consumption optimisation in FPGA-based heterogeneous HPC systems" [16] provides a comprehensive investigation into optimising energy efficiency in heterogeneous High-Performance Computing (HPC) systems, with a focus on integrating Field-Programmable Gate Arrays (FPGAs) into existing architectures. The authors, Oleksandr V. Hryshchuk and Sergiy P. Zagorodnyuk from Taras Shevchenko National University of Kyiv, Ukraine, delve into the parametrisation, modelling, and optimisation techniques necessary for sustainable HPC practices.

The article begins by outlining the growing concern over the escalating energy consumption of HPC systems, highlighting the need for effective optimisation strategies to address sustainability and operational costs. It characterises the heterogeneity within modern HPC environments, incorporating diverse hardware components such as CPUs, GPUs, FPGAs, and accelerators.

The research delves into modelling techniques, leveraging heuristic methods and statistical approaches to construct accurate predictive models for energy consumption. Additionally,

Figure 11: Presentation of paper [19].

integrating dynamic power management strategies, such as Dynamic Voltage and Frequency Scaling (DVFS) and task scheduling, is explored to optimise energy usage without compromising performance.

The authors provide a theoretical framework for energy optimisation in heterogeneous HPC systems, discussing optimisation problem definitions for task scheduling and outlining optimisation criteria. They compare cluster architectures, focusing on homogeneous Massive Parallel Processor (MPP) systems and heterogeneous systems combining CPUs, GPUs, and FPGAs. The article highlights the emerging field of FPGA-based HPC systems and identifies a research gap in energy optimisation for these systems.

In conclusion, the article emphasises the need for further research and development of energy optimisation techniques tailored to FPGA-based heterogeneous HPC systems. It suggests that future work should amplify existing methods, including heuristic solutions for power consumption planning in FPGA-coupled architectures.

The article "Conference platform metadata and functions: existing platforms analysis and ontology-based approach" by Shapovalov and Shapovalov [17] provides a comprehensive analysis of existing conference management platforms and proposes an ontology-based approach to enhance the structure and functionality of such systems. The review begins by highlighting the rise of predatory conferences and the need for robust platforms to ensure the quality and integrity of scholarly events.

The authors analyse six well-known conference platforms, categorising them into informational-oriented and process-oriented systems. Each platform is detailed, emphasising its unique features and focus areas. The authors identify standard fields and functionalities across these platforms through data collection and processing, revealing insights into user priorities and platform capabilities.

Key findings include the prevalence of search functionality as the most critical feature, followed by peer reviewing, registration, submission, and publication of conference materials. Additionally, identifiers such as DOI and subject-specific databases like DBLP are highlighted for their role in the accurate cataloguing and citation of academic work.

The article proposes an ontology-based approach to organise conference data, leveraging systems like CIT Polyhedron to provide flexible data structures. This approach is a solution to counteract predatory conferences by promoting healthy competition and ensuring structured data entry.

5.4. Computer applications

The article "Dynamic system analysis using telemetry" by Talaver and Vakaliuk [21] provides a comprehensive exploration of dynamic system analysis using telemetry, focusing on detecting harmful architectural practices and anomalous events in distributed information systems. It begins by highlighting the increasing complexity introduced by distributed architectures like microservices, necessitating advanced monitoring and analysis tools to ensure system performance and reliability.

The theoretical background section effectively contextualises the study within the evolution of system observability, particularly emphasising the role of telemetry in providing a holistic view of system behaviour. The discussion on the OpenTelemetry standard and its role in

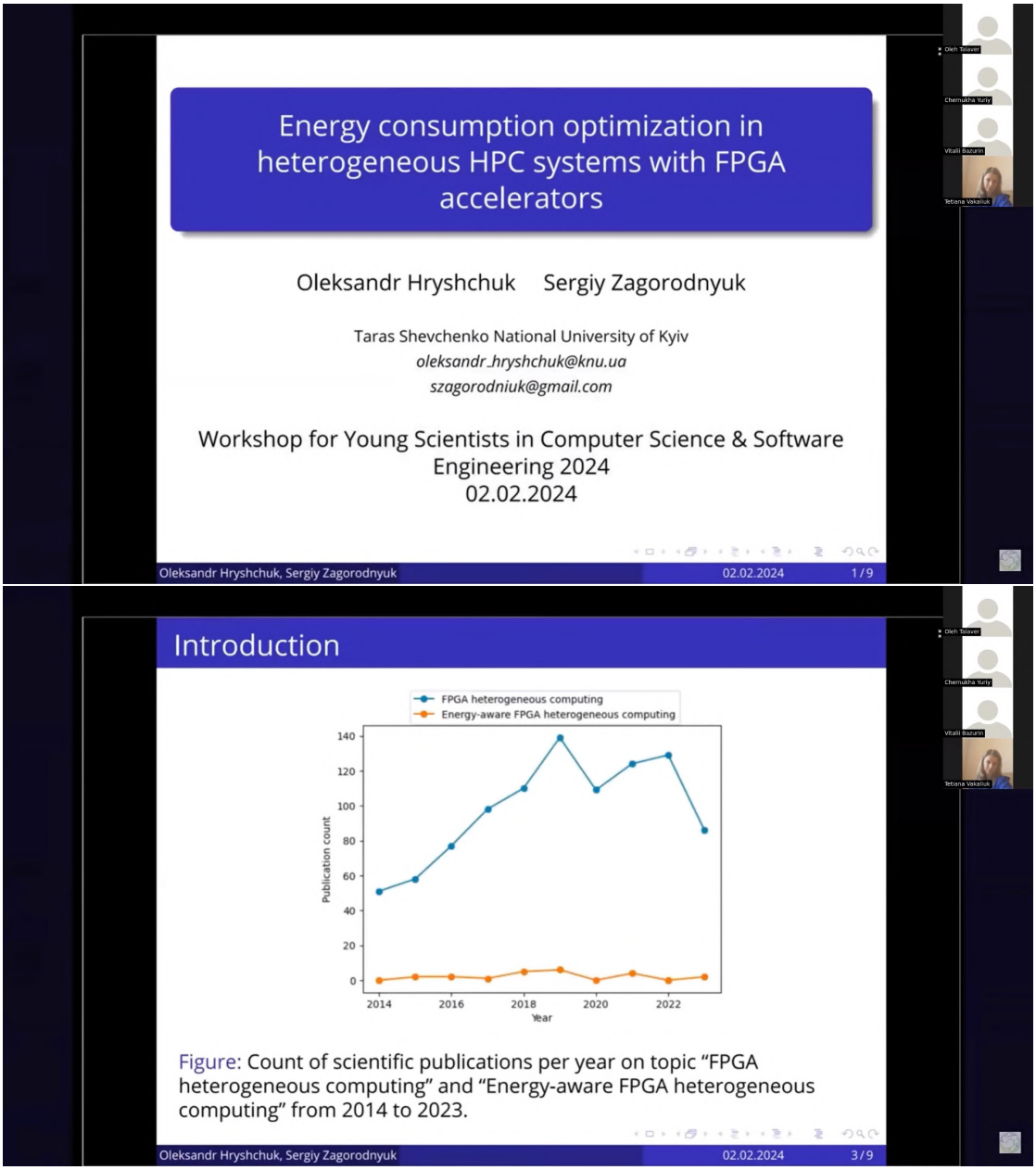


Figure 12: Presentation of paper [16].

unifying telemetry collection and analysis is informative, highlighting its significance in modern monitoring practices.

The methods section details the approach, covering data collection, storage, and analysis. The choice of Neo4j as the graph database management system for storing system models is justified and the integration of telemetry data into the graph structure is well-explained. Additionally, the

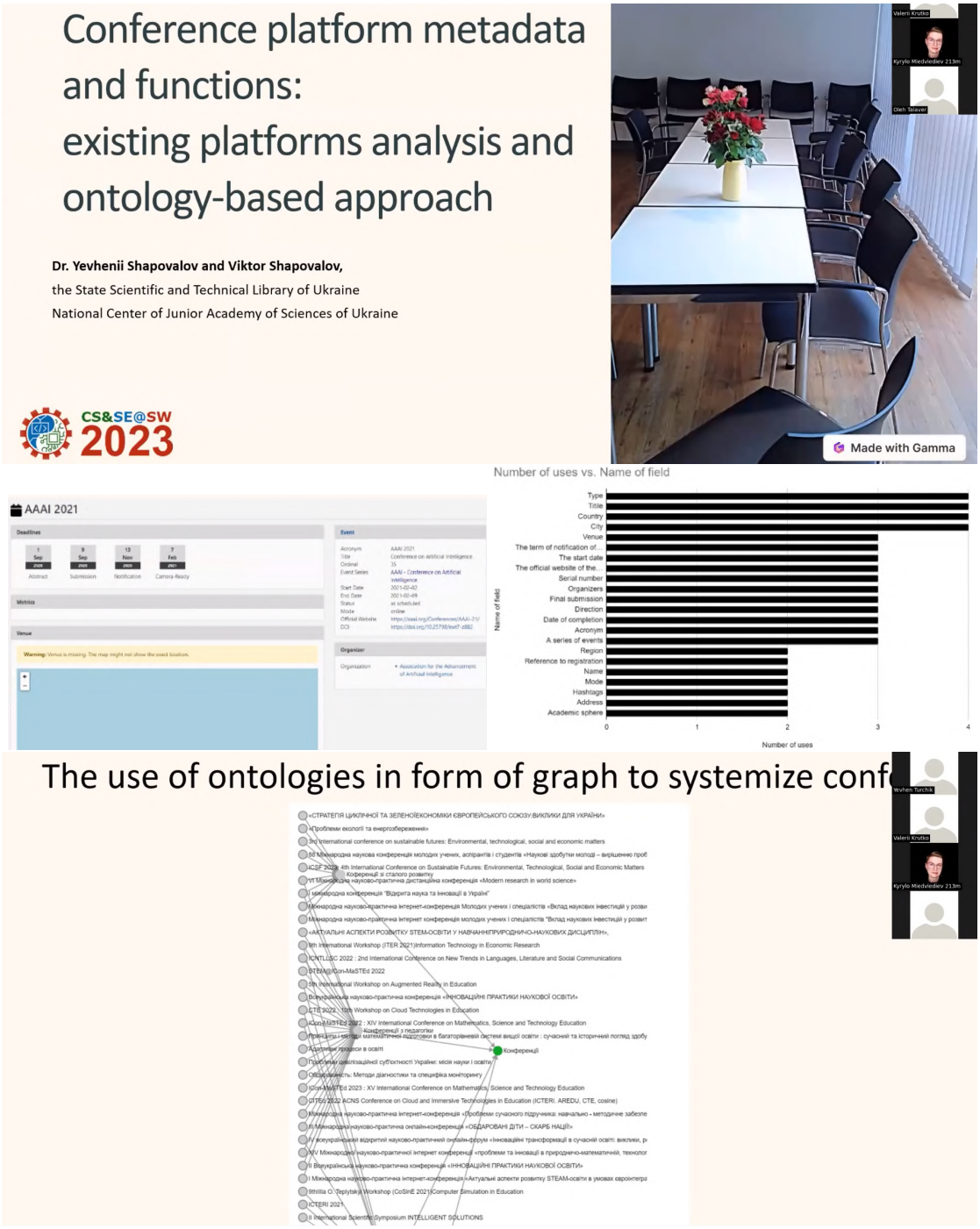


Figure 13: Presentation of paper [17].

explanation of anomaly detection using the PCA algorithm is clear and insightful, showcasing how statistical methods can be leveraged for identifying system irregularities.

Results are presented effectively through visualisations generated from the Neo4j database, demonstrating the practical application of the proposed methodology. Using Neo4j Bloom to visualise service dependencies and anomalies adds clarity to the analysis, making it easier to identify potential areas of improvement in system architecture.

The discussion section provides valuable insights into the advantages of dynamic analysis over static approaches and the potential for further development in telemetry-based analysis. The comparison with existing approaches, such as New Relic, highlights the strengths of the proposed method while acknowledging areas for future enhancement.

The article "Development of a modified genetic method for automatic university scheduling" by Fedorchenko et al. [22] from the National University "Zaporizhzhia Polytechnic" in Ukraine addresses the challenging task of optimising university class schedules, crucial for adequate time and resource management in higher education.

The authors propose a modified genetic algorithm for university scheduling, aiming to minimise conflicts and time intervals between classes while considering recommendations for time and place. The paper outlines the development of sequential and parallel methods for scheduling based on genetic search, incorporating adapted initialisation, crossover, and selection operators.

A comparative analysis between classical and modified genetic algorithms is presented, confirming the efficiency of the proposed approach. The modified algorithm is also compared with different operators and parameters to determine optimal conditions. The results demonstrate effective methods for improving schedule quality and optimising the educational process.

The article provides a detailed literature review, problem statement, and mathematical model development for university scheduling optimisation. It describes the software implementation of the proposed modification and conducts experiments to evaluate its performance.

The article "Predictive machine learning of soybean oil epoxidising reactions using artificial neural networks" by Sus et al. [20] presents an innovative approach to optimising the epoxidation process of soybean oil through the utilisation of artificial neural networks (ANNs). The study employs experimental data to construct a training dataset for the ANN, which then facilitates the optimisation of epoxy curing reaction parameters, monitors its evolution, and refines the epoxy product synthesis process.

The authors discuss the broad applicability of neural networks across various scientific and technological domains, highlighting their importance in predicting outcomes, selecting optimal conditions, and assessing quantities in chemical and biological processes. They emphasise the significance of green chemistry and the growing importance of soybean oil epoxidation in various industrial applications.

The experimental setup involves the epoxidation of soybean oil using a specific hydrogen peroxide system, acetic anhydride, and a catalyst. The study explores various parameters such as concentration of reactants, catalyst amount, temperature, and reaction time. A neural network model is then trained using this experimental data to predict the outcomes of the epoxidation process.

Results indicate that the neural network accurately predicts the epoxy and iodine numbers, crucial indicators of the quality of epoxidised oils, based on the input parameters. The authors

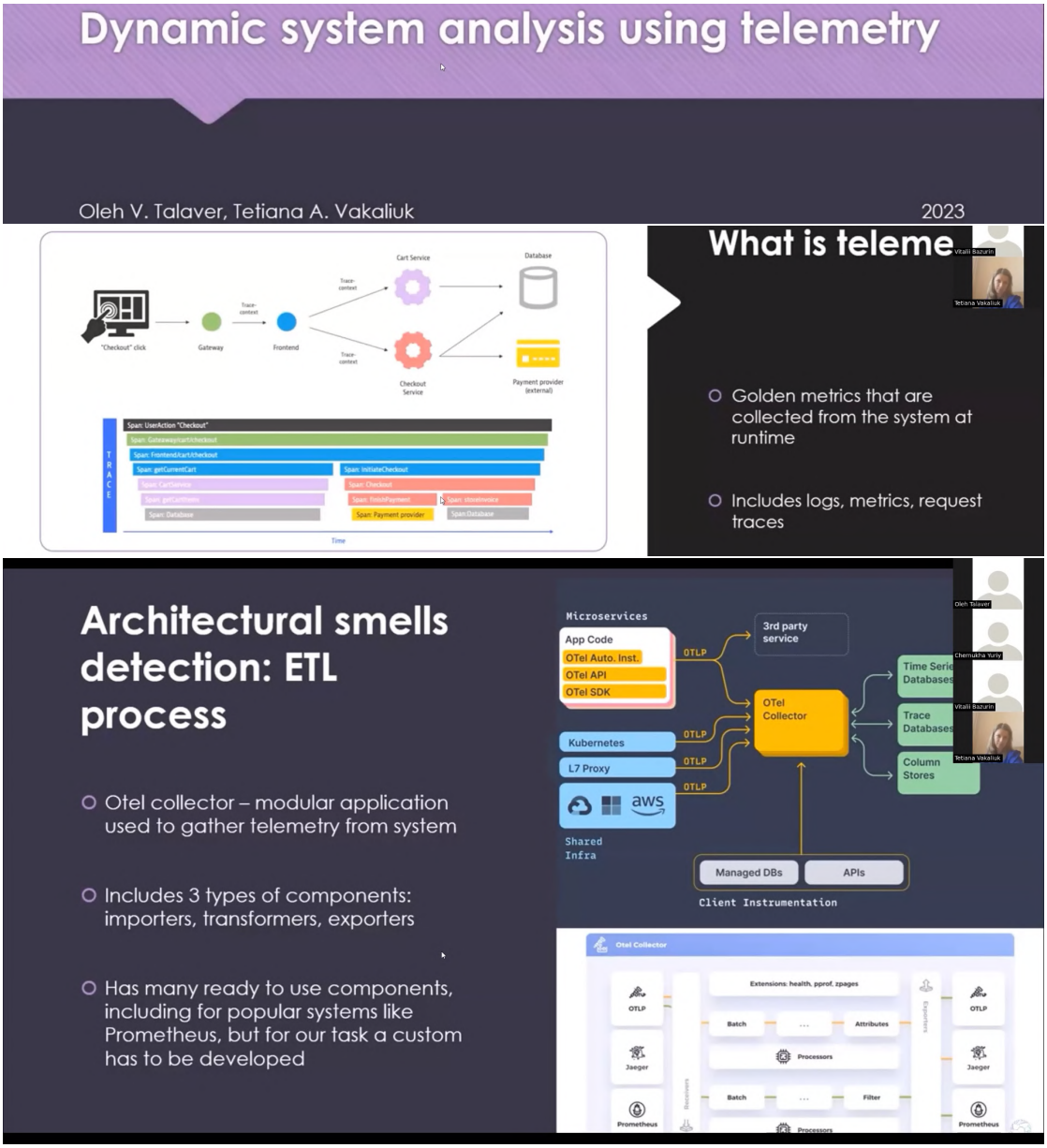


Figure 14: Presentation of paper [21].

demonstrate the network's ability to interpolate experimental data to generate comprehensive dependency graphs, even beyond the scope of available experimental data.

Moreover, the study identifies optimal conditions for maximising the epoxy number and minimising the iodine number during the epoxidation process, showcasing the practical utility of the neural network in process optimisation.

In conclusion, the article presents a robust methodology for optimising soybean oil epoxida-

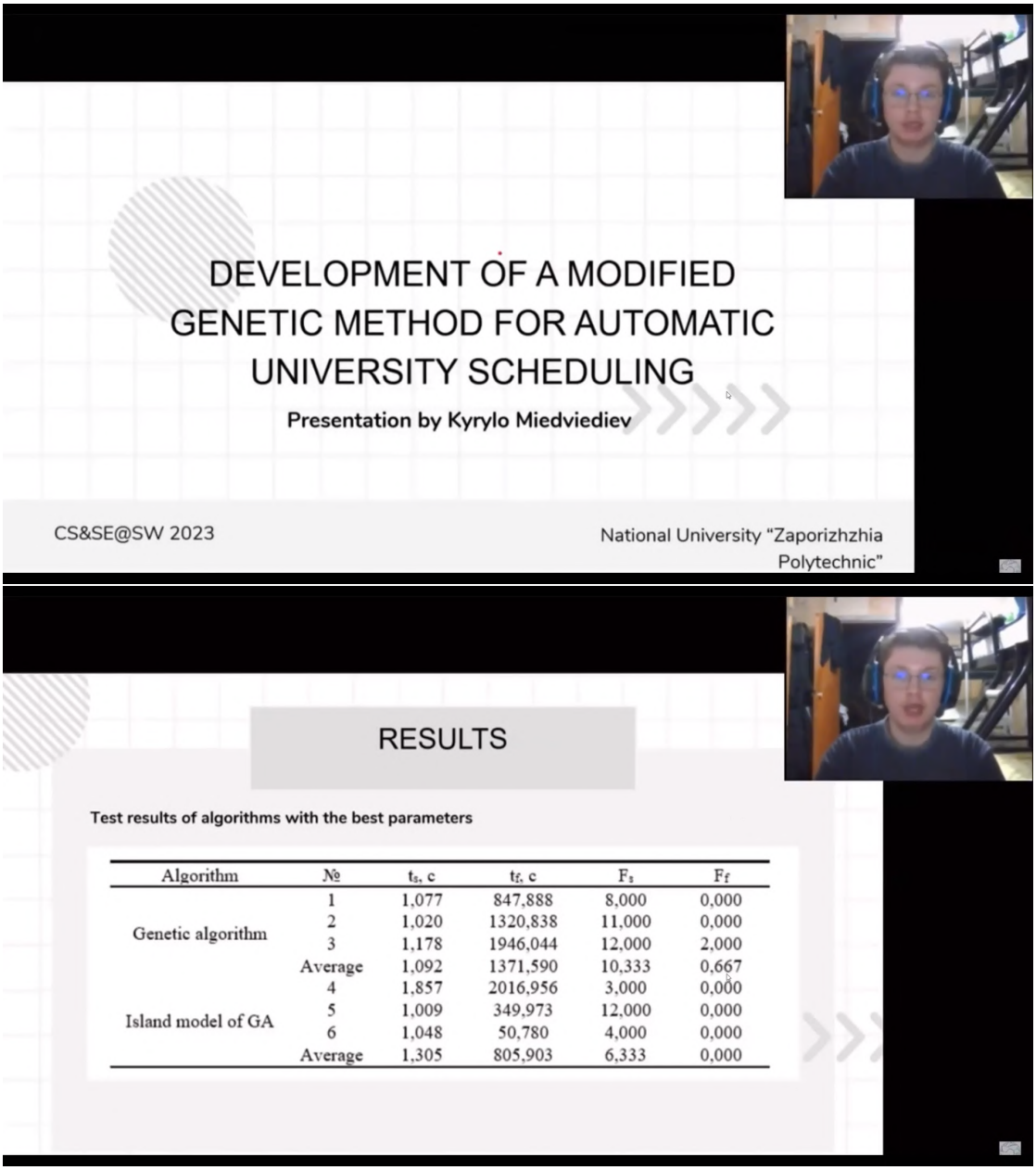


Figure 15: Presentation of paper [22].

tion using predictive machine learning, offering insights into reaction parameters and paving the way for further advancements in the field. The approach holds promise for soybean oil and other vegetable oils, expanding its applicability across various industrial processes. Overall, the article provides valuable contributions to both the fields of chemical engineering and machine learning.

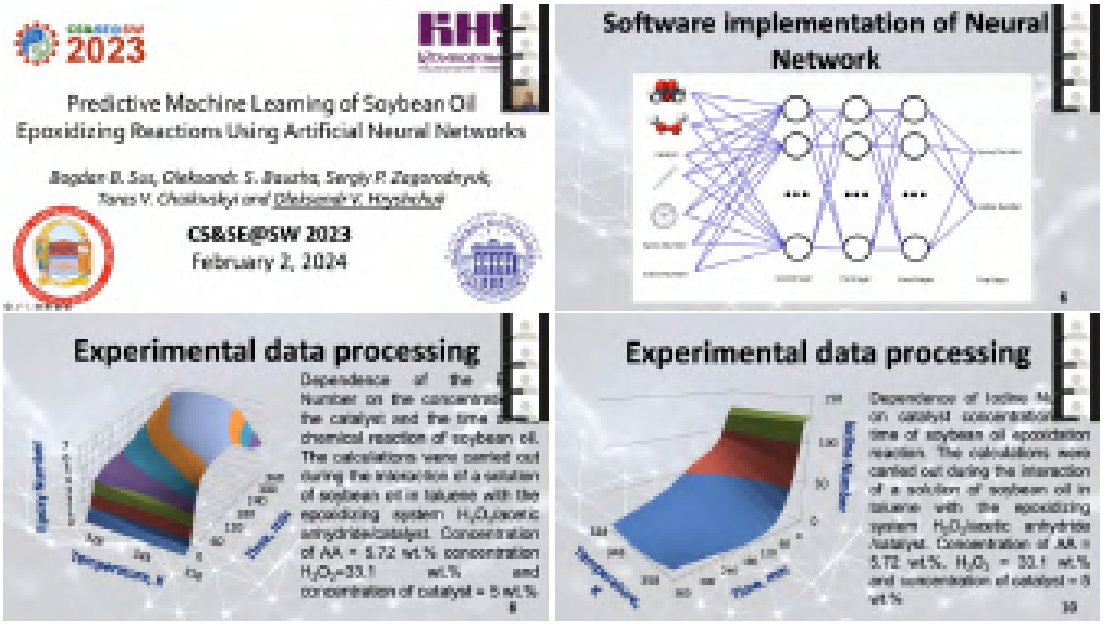


Figure 16: Presentation of paper [20].

The article "Software development of thermal resistance calculator for thermal insulation parameters determines dielectric building structures" by Bazurin et al. [10] presents a detailed review of the software development of a thermal resistance calculator named "ThermoResist" for determining the parameters of thermal insulation in dielectric building structures. The calculator is designed to calculate thermal resistance according to the State Building Regulations of Ukraine, assuming that the contributions of different thermal resistance mechanisms are additive.

The authors provide an in-depth discussion of the computational method involved, which includes formulas and theoretical background related to thermal conductivity and thermal resistance in dielectric materials. They emphasise the importance of accurate prediction of thermal conductivity in construction, particularly in rebuilding efforts post-war in Ukraine. The article also compares existing thermal resistance calculators and identifies their limitations, leading to the development of a specialised tool like "ThermoResist".

The functionalities of "ThermoResist" are described in detail, including modules for calculating the thermal resistance of walls, windows, attic floors, and roof overlaps. The calculator's interface is intuitive, allowing users to easily input relevant data and obtain thermal resistance calculations. The article also provides a class diagram of the program's structure and discusses the choice of programming language (C♯) and development environment (Microsoft Visual Studio 2022).

In conclusion, the authors highlight the significance of digitalisation in society and the importance of tools like "ThermoResist" in the construction industry. They emphasise that the calculator adheres to State Building Regulations and can be beneficial for both educational purposes and practical applications by civil engineers.

Figure 17: Presentation of paper [10].

The article “Using the Three.js library to develop remote physical laboratory to investigate diffraction” [12] presents a detailed examination of the process involved in designing and developing a virtual physics laboratory focused on studying the diffraction effect. Authored by Pavlo I. Chopyk, Vasyl P. Oleksiuk, and Oleksandr P. Chukhrai from Ternopil Volodymyr Hnatiuk National Pedagogical University in Ukraine, the article addresses the requirements,

framework selection, design, and implementation of the virtual laboratory.

The authors begin by outlining the importance of laboratory experiments in physics education, highlighting their role in facilitating understanding, skill development, and critical thinking. They also acknowledge the increasing prevalence of remote training and the need for virtual laboratories to supplement traditional methods, mainly when practical experience is limited or hazardous.

The article systematically discusses the criteria for selecting the appropriate development tools, focusing on 3D graphics libraries. After conducting a comparative analysis, the authors choose the Three.js library for its performance, ease of use, flexibility, feature set, and compatibility. They then describe the stages of designing and developing the virtual laboratory, including formulating the physical problem, selecting tools, creating the laboratory model, and implementing and testing.

Detailed explanations accompanied by code snippets illustrate the creation of the virtual laboratory components, such as scene objects, lighting, cameras, and interactive controls. The authors emphasise the importance of accurately simulating the diffraction phenomenon and providing students with tools for measurement and analysis, ensuring a realistic and educational experience.

The virtual laboratory developed using Three.js allows students to observe diffraction patterns, measure distances, and calculate wavelengths, mimicking real-world experimental setups. The article discusses integrating features such as dynamic screens, rulers, and colour filters, providing students with a comprehensive learning environment.

Finally, the authors compare the results obtained from the virtual laboratory with those from natural experiments, demonstrating the accuracy and effectiveness of the virtual simulation. They also acknowledge limitations such as hardware requirements and outline future research directions, including collaboration features and integration with learning management systems.

6. CS&SE@SW 2023: Conclusion and outlook

The 6th Workshop for Young Scientists in Computer Science & Software Engineering (CS&SE@SW 2023) has once again demonstrated its commitment to fostering the growth of emerging researchers and providing a platform for exchanging innovative ideas and early research findings. The diverse range of papers presented at this year's workshop showcases the breadth and depth of the research undertaken by young scientists, covering various topics within computer science and software engineering.

The vision of CS&SE@SW 2023 has been to create an expert environment where young researchers can present and discuss their cutting-edge work, receive valuable feedback from peers and experienced academics, and establish collaborations that transcend geographical boundaries. The workshop has proven to be a nurturing ground for developing research skills, critical thinking, and the dissemination of knowledge.

The proceedings of CS&SE@SW 2023 reflect the multifaceted nature of the challenges and opportunities that lie ahead in the rapidly evolving fields of computer science and software engineering. From exploring emerging technologies such as telemetry, graph theory, and machine learning for optimising distributed systems and detecting anomalies to investigating

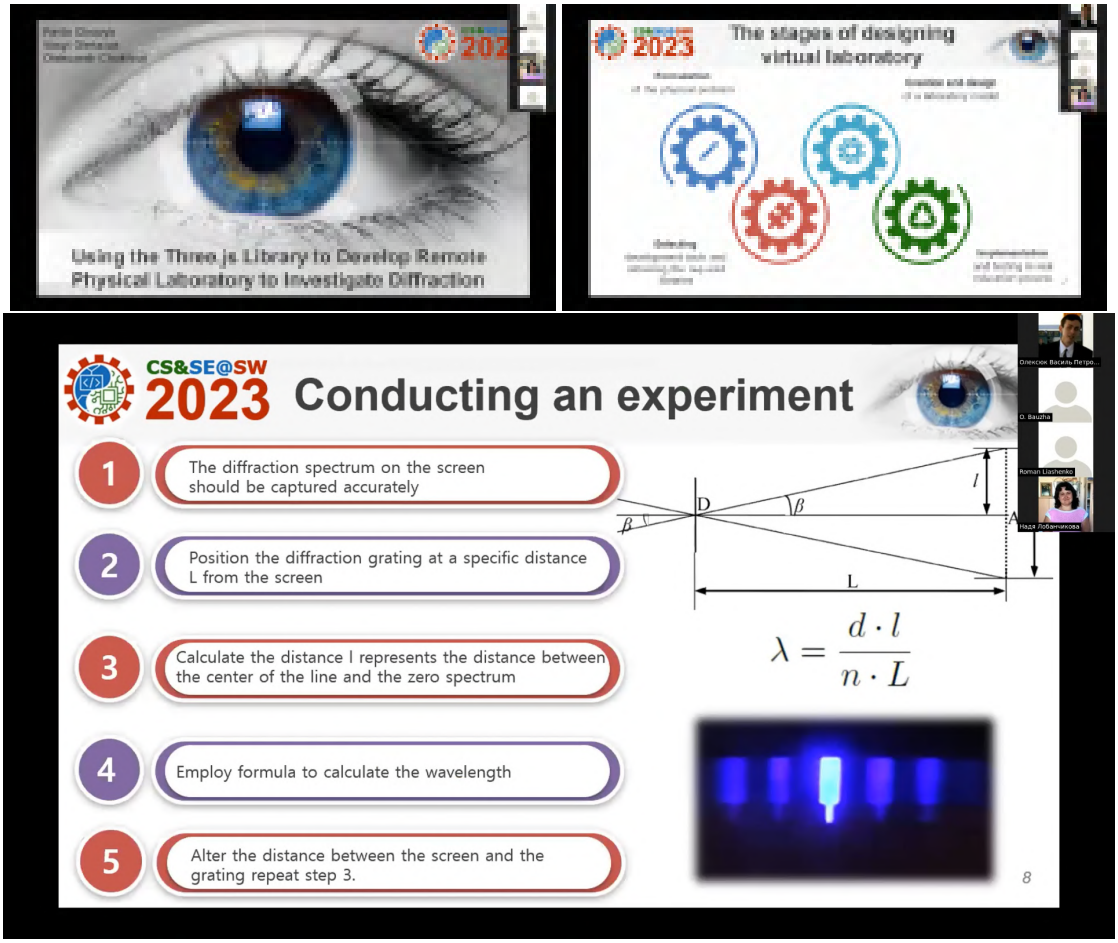


Figure 18: Presentation of paper [12].

acoustic surveillance techniques for UAV detection and employing genetic algorithms for university scheduling, the contributions showcased in this volume demonstrate the remarkable diversity and ingenuity of the research community.

Furthermore, the workshop has delved into software reliability assessment, user experience research in cross-platform application development, virtual physics laboratories, and Python learning game applications, underscoring the importance of interdisciplinary approaches and the fusion of theory and practice.

Looking ahead, CS&SE@SW 2023 has laid the foundation for future collaborations, fostering a spirit of curiosity, innovation, and critical inquiry among young scientists. The insights and findings presented during the workshop will undoubtedly catalyse further exploration, igniting new avenues of research and propelling the fields of computer science and software engineering towards new horizons.

As we conclude this successful edition of the workshop, we extend our gratitude to all the authors, delegates, program committee members, and peer reviewers who have contributed

to its success. Their invaluable efforts and commitment have ensured the high quality and relevance of the presented work, further elevating the standards of academic excellence.

We look forward to the next instalment of CS&SE@SW, scheduled for December 20, 2024, in Kryvyi Rih, Ukraine. This future gathering promises to be an even more enriching and thought-provoking experience, where emerging talents will converge to share their latest discoveries, engage in stimulating discussions, and forge lasting connections that will shape the future of these dynamic and ever-evolving fields.

References

[1] A. E. Kiv, S. O. Semerikov, V. N. Soloviev, A. M. Striuk, First Student Workshop on Computer Science & Software Engineering, CEUR Workshop Proceedings 2292 (2018) 1–10. URL: http://ceur-ws.org/Vol-2292/paper00.pdf.

[2] A. E. Kiv, S. O. Semerikov, V. N. Soloviev, A. M. Striuk, Second Student Workshop on Computer Science & Software Engineering, CEUR Workshop Proceedings 2546 (2019) 1–20. URL: http://ceur-ws.org/Vol-2546/paper00.pdf.

[3] A. E. Kiv, S. O. Semerikov, V. N. Soloviev, A. M. Striuk, 3rd Workshop for Young Scientists in Computer Science & Software Engineering, CEUR Workshop Proceedings 2832 (2020) 1–10. URL: http://ceur-ws.org/Vol-2832/paper00.pdf.

[4] A. E. Kiv, S. O. Semerikov, V. N. Soloviev, A. M. Striuk, 4th Workshop for Young Scientists in Computer Science & Software Engineering, CEUR Workshop Proceedings 3077 (2022) i–xxxv. URL: https://ceur-ws.org/Vol-3077/intro.pdf.

[5] S. O. Semerikov, A. M. Striuk, Embracing Emerging Technologies: Insights from the 6th Workshop for Young Scientists in Computer Science & Software Engineering, CEUR Workshop Proceedings (2024) 1–36.

[6] Y. O. Chernukha, O. V. Klochko, T. P. Zuziak, Methodology of implementation of modern information systems at commercial enterprises, CEUR Workshop Proceedings (2024) 48–62.

[7] V. P. Oleksiuk, D. V. Verbovetskyi, I. A. Hrytsai, Design and development of a game application for learning Python, CEUR Workshop Proceedings (2024) 111–124.

[8] M. Y. Salohub, O. H. Rybalchenko, S. V. Bilashenko, Designing a cross-platform user-friendly transport company application, CEUR Workshop Proceedings (2024) 75–85.

[9] Y. L. Turchyk, M. V. Puzino, O. H. Rybalchenko, S. V. Bilashenko, Research of the route planning algorithms on the example of a drone delivery system software development, CEUR Workshop Proceedings (2024) 86–100.

[10] V. M. Bazurin, O. I. Pursky, Y. M. Karpenko, T. V. Pidhorna, A. I. Nechepourenko, Software development of thermal resistance calculator for thermal insulation parameters determines dielectric building structures, CEUR Workshop Proceedings (2024) 237–245.

[11] I. V. Krasnokutska, O. S. Krasnokutskyi, Implementing E2E tests with Cypress and Page Object Model: evolution of approaches, CEUR Workshop Proceedings (2024) 101–110.

[12] P. I. Chopyk, V. P. Oleksiuk, O. P. Chukhrai, Using the Three.js library to develop remote physical laboratory to investigate diffraction, CEUR Workshop Proceedings (2024) 246–259.

[13] N. Rudnichenko, V. Vychuzhanin, T. Otradskya, I. Petrov, Information system module for analysis viral infections data based on machine learning, CEUR Workshop Proceedings (2024) 63–74.
[14] V. Krutko, I. Spivak, S. Krepych, An approach to assessing the reliability of software systems based on a graph model of method dependence, CEUR Workshop Proceedings (2024) 37–47.
[15] O. V. Solomentsev, M. Y. Zaliskyi, D. I. Bakhtiiarov, B. S. Chumachenko, Data processing method for multimodal distribution parameters estimation, CEUR Workshop Proceedings (2024) 144–154.
[16] O. V. Hryshchuk, S. P. Zagorodnyuk, Modern methods of energy consumption optimization in FPGA-based heterogeneous HPC systems, CEUR Workshop Proceedings (2024) 167–176.
[17] Y. B. Shapovalov, V. B. Shapovalov, Conference platform metadata and functions: existing platforms analysis and ontology-based approach, CEUR Workshop Proceedings (2024) 177–192.
[18] O. Y. Lavrynenko, D. I. Bakhtiiarov, B. S. Chumachenko, O. G. Holubnychyi, G. F. Konakhovych, V. V. Antonov, Application of Daubechies wavelet analysis in problems of acoustic detection of UAVs, CEUR Workshop Proceedings (2024) 125–143.
[19] I. V. Ponomarenko, V. M. Pavlenko, O. B. Morhulets, D. V. Ponomarenko, N. M. Ukhnal, Application of artificial intelligence in digital marketing, CEUR Workshop Proceedings (2024) 155–166.
[20] B. B. Sus, O. S. Bauzha, S. P. Zagorodnyuk, T. V. Chaikivskyi, , O. V. Hryshchuk, Predictive machine learning of soybean oil epoxidizing reactions using artificial neural networks, CEUR Workshop Proceedings (2024) 223–236.
[21] O. V. Talaver, T. A. Vakaliuk, Dynamic system analysis using telemetry, CEUR Workshop Proceedings (2024) 193–209.
[22] I. Fedorchenko, A. Oliinyk, T. Zaiko, K. Miedviediev, Y. Fedorchenko, M. Khokhlov, Development of a modified genetic method for automatic university scheduling, CEUR Workshop Proceedings (2024) 210–222.
[23] N. Cavus, M. M. Al-Momani, Mobile system for flexible education, Procedia Computer Science 3 (2011) 1475–1479. doi:10.1016/j.procs.2011.01.034, world Conference on Information Technology.
[24] A. B. Mbombo, N. Cavus, Smart University: A University In the Technological Age, TEM Journal (2021) 13–17. doi:10.18421/tem101-02.
[25] D. Budgen, J. Bailey, M. Turner, B. Kitchenham, P. Brereton, S. Charters, Cross-domain investigation of empirical practices, IET Software 3 (2009) 410–421(11). URL: https://digital-library.theiet.org/content/journals/10.1049/iet-sen.2008.0106.
[26] D. Budgen, B. Kitchenham, S. Charters, S. Gibbs, A. Pohthong, J. Keung, P. Brereton, Lessons from Conducting a Distributed Quasi-experiment, in: 2013 ACM / IEEE International Symposium on Empirical Software Engineering and Measurement, 2013, pp. 143–152. doi:10.1109/ESEM.2013.12.
[27] A. Kertész, P. Kacsuk, A Taxonomy of Grid Resource Brokers, in: P. Kacsuk, T. Fahringer, Z. Németh (Eds.), Distributed and Parallel Systems, Springer US, Boston, MA, 2007, pp. 201–210. doi:10.1007/978-0-387-69858-8_20.

[28] B. Mishra, B. Mishra, A. Kertesz, Stress-Testing MQTT Brokers: A Comparative Analysis of Performance Measurements, Energies 14 (2021) 5817. doi:10.3390/en14185817.

[29] J. Suryadevara, B. Sunil, N. K. Suryadevara, Secured multimedia authentication system for wireless sensor network data related to internet of things, in: Seventh International Conference on Sensing Technology, ICST 2013, Wellington, New Zealand, December 3-5, 2013, IEEE, 2013, pp. 109–115. URL: https://doi.org/10.1109/ICSensT.2013.6727625. doi:10.1109/ICSENST.2013.6727625.

[30] N. K. Survadevara, S. C. Mukhopadhyay, R. K. Rayudu, Applying SARIMA time series to forecast sleeping activity for wellness model of elderly monitoring in smart home, in: 2012 Sixth International Conference on Sensing Technology (ICST), 2012, pp. 157–162. doi:10.1109/ICSensT.2012.6461661.

[31] M. I. Nadeem, K. Ahmed, D. Li, Z. Zheng, H. K. Alkahtani, S. M. Mostafa, O. Mamyrbayev, H. Abdel Hameed, EFND: A Semantic, Visual, and Socially Augmented Deep Framework for Extreme Fake News Detection, Sustainability 15 (2023) 133. doi:10.3390/su15010133.

[32] A. Yeshmukhametov, M. Kalimoldayev, O. Mamyrbayev, Y. Amirgaliev, Design and kinematics of serial/parallel hybrid robot, in: 2017 3rd International Conference on Control, Automation and Robotics (ICCAR), 2017, pp. 162–165. doi:10.1109/ICCAR.2017.7942679.

[33] J. Bae, B. Moon, Time synchronization with fast asynchronous diffusion in wireless sensor network, in: 2009 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery, 2009, pp. 82–85. doi:10.1109/CYBERC.2009.5342158.

[34] H. Lee, B. Moon, A. H. Aghvami, Enhanced SIP for Reducing IMS Delay under WiFi-to-UMTS Handover Scenario, in: 2008 The Second International Conference on Next Generation Mobile Applications, Services, and Technologies, 2008, pp. 640–645. doi:10.1109/NGMAST.2008.63.

[35] J. Wan, C. A. Byrne, M. J. O'Grady, G. M. P. O'Hare, Managing Wandering Risk in People With Dementia, IEEE Transactions on Human-Machine Systems 45 (2015) 819–823. doi:10.1109/THMS.2015.2453421.

[36] C. Muldoon, G. M. P. O'Hare, M. J. O'Grady, R. Tynan, Agent Migration and Communication in WSNs, in: 2008 Ninth International Conference on Parallel and Distributed Computing, Applications and Technologies, 2008, pp. 425–430. doi:10.1109/PDCAT.2008.58.

[37] J. Morajda, G. Paliwoda-Pekosz, An Enhancement of Kohonen Neural Networks for Predictive Analytics: Self-Organizing Prediction Maps, in: B. B. Anderson, J. Thatcher, R. D. Meservy, K. Chudoba, K. J. Fadel, S. Brown (Eds.), 26th Americas Conference on Information Systems, AMCIS 2020, Virtual Conference, August 15-17, 2020, Association for Information Systems, 2020. URL: https://aisel.aisnet.org/amcis2020/ai_semantic_for_intelligent_info_systems/ai_semantic_for_intelligent_info_systems/6.

[38] P. Lula, G. Paliwoda-Pundefinedkosz, An ontology-based cluster analysis framework, in: Proceedings of the First International Workshop on Ontology-Supported Business Intelligence, OBI '08, Association for Computing Machinery, New York, NY, USA, 2008. doi:10.1145/1452567.1452574.

[39] E. Serral, P. Valderas, V. Pelechano, Addressing the evolution of automated user behaviour

patterns by runtime model interpretation, Software & Systems Modeling 14 (2015) 1387–1420. doi:10.1007/s10270-013-0371-3.

[40] E. Serral, P. Valderas, V. Pelechano, A Model Driven Development Method for Developing Context-Aware Pervasive Systems, in: F. E. Sandnes, Y. Zhang, C. Rong, L. T. Yang, J. Ma (Eds.), Ubiquitous Intelligence and Computing, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008, pp. 662–676. doi:10.1007/978-3-540-69293-5_52.

[41] Y. Romanenkov, V. Pasichnyk, N. Veretennikova, M. Nazaruk, A. Leheza, Information and Technological Support for the Processes of Prognostic Modeling of Regional Labor Markets, CEUR Workshop Proceedings 2386 (2019) 24–34. URL: https://ceur-ws.org/Vol-2386/paper3.pdf.

[42] N. Veretennikova, N. Kunanets, Recommendation Systems as an Information and Technology Tool for Virtual Research Teams, in: N. Shakhovska, V. Stepashko (Eds.), Advances in Intelligent Systems and Computing II, Springer International Publishing, Cham, 2018, pp. 577–587. doi:10.1007/978-3-319-70581-1_40.

[43] M. Dong, L. Yao, X. Wang, B. Benatallah, Q. Z. Sheng, H. Huang, DUAL: A Deep Unified Attention Model with Latent Relation Representations for Fake News Detection, in: H. Hacid, W. Cellary, H. Wang, H.-Y. Paik, R. Zhou (Eds.), Web Information Systems Engineering – WISE 2018, Springer International Publishing, Cham, 2018, pp. 199–209. doi:10.1007/978-3-030-02922-7_14.

[44] K. Chen, L. Yao, X. Wang, D. Zhang, T. Gu, Z. Yu, Z. Yang, Interpretable Parallel Recurrent Neural Networks with Convolutional Attentions for Multi-Modality Activity Modeling, in: 2018 International Joint Conference on Neural Networks (IJCNN), 2018, pp. 1–8. doi:10.1109/IJCNN.2018.8489767.

[45] A. Zunino, M. Campo, Chronos: A multi-agent system for distributed automatic meeting scheduling, Expert Systems with Applications 36 (2009) 7011–7018. doi:10.1016/j.eswa.2008.08.024.

[46] A. De Renzis, M. Garriga, A. Flores, A. Cechich, A. Zunino, Case-based Reasoning for Web Service Discovery and Selection, Electronic Notes in Theoretical Computer Science 321 (2016) 89–112. doi:10.1016/j.entcs.2016.02.006, cLEI 2015, the XLI Latin American Computing Conference.

[47] B. Schooley, N. Hikmet, E. Atilgan, Health IT Maturity and Hospital Quality: Effects of PACS Automation and Integration Levels on U.S. Hospital Performance, in: 2016 International Conference on Computational Science and Computational Intelligence (CSCI), 2016, pp. 45–50. doi:10.1109/CSCI.2016.0016.

[48] E. Atilgan, I. Ozcelik, E. N. Yolacan, MQTT Security at a Glance, in: 2021 International Conference on Information Security and Cryptology (ISCTURKEY), 2021, pp. 138–142. doi:10.1109/ISCTURKEY53027.2021.9654337.

[49] I. Krak, O. Barmak, E. Manziuk, A. Kulias, Data Classification Based on the Features Reduction and Piecewise Linear Separation, in: P. Vasant, I. Zelinka, G.-W. Weber (Eds.), Intelligent Computing and Optimization, Springer International Publishing, Cham, 2020, pp. 282–289. doi:10.1007/978-3-030-33585-4_28.

[50] Y. Krak, O. Barmak, O. Mazurets, The practice implementation of the information technology for automated definition of semantic terms sets in the content of educational materials, CEUR Workshop Proceedings 2139 (2018) 245–254. URL: http://ceur-ws.org/

Vol-2139/245-254.pdf.
[51] K. M. Caramancion, The Relation Between Time of the Day and Misinformation Vulnerability: A Multivariate Approach, in: 2021 IEEE 16th International Conference on Computer Sciences and Information Technologies (CSIT), volume 1, 2021, pp. 150–153. doi:10.1109/CSIT52700.2021.9648654.
[52] K. M. Caramancion, Textual vs. Visual Fake News: A Deception Showdown, in: 2021 IEEE International Conference on Cloud Computing in Emerging Markets (CCEM), 2021, pp. 31–35. doi:10.1109/CCEM53267.2021.00015.
[53] P. Hryhoruk, N. Khrushch, S. a. Grygoruk, Using Multidimensional Scaling for Assessment Economic Development of Regions, International journal of industrial Engineering & Production Research 31 (2020). doi:10.22068/ijiepr.31.4.597.
[54] P. Hryhoruk, N. Khrushch, S. Grygoruk, K. Gorbatiuk, L. Prystupa, Assessing the Impact of COVID-19 Pandemic on the Regions' Socio-Economic Development: The Case of Ukraine, European Journal of Sustainable Development 10 (2021) 63. doi:10.14207/ejsd.2021.v10n1p63.
[55] V. N. Kukharenko, A. P. Fedosova, A. G. Kolgatin, V. G. Dosov, Studying the processes in the xenon heat exchanger-freezer, Khimicheskoe I Neftegazovoe Mashinostroenie (1992) 19–21.
[56] L. Bilousova, O. Kolgatin, L. Kolgatina, Pedagogical Diagnostics with Use of Computer Technologies, CEUR Workshop Proceedings 1000 (2013) 209–220. URL: https://ceur-ws.org/Vol-1000/ICTERI-2013-p-209-220.pdf.
[57] A. V. Riabko, T. A. Vakaliuk, O. V. Zaika, R. P. Kukharchuk, V. V. Kontsedailo, Chatbot algorithm for solving physics problems, CEUR Workshop Proceedings 3553 (2023) 75–92. URL: https://ceur-ws.org/Vol-3553/paper5.pdf.
[58] A. V. Riabko, T. A. Vakaliuk, O. V. Zaika, R. P. Kukharchuk, V. V. Kontsedailo, Cluster fault tolerance model with migration of virtual machines, CEUR Workshop Proceedings 3374 (2023) 23–40. URL: https://ceur-ws.org/Vol-3374/paper02.pdf.
[59] A. Hrechuk, V. Bushlya, J.-E. Ståhl, V. Kryzhanivskyy, Novel metric "Implenarity" for characterization of shape and defectiveness: The case of CFRP hole quality, Composite Structures 265 (2021) 113722. doi:10.1016/j.compstruct.2021.113722.
[60] M. Moreno, J. M. Andersson, R. M'Saoubi, V. Kryzhanivskyy, M. P. Johansson-Jöesaar, L. J. S. Johnson, M. Odén, L. Rogström, Adhesive wear of tialn coatings during low speed turning of stainless steel 316l, Wear 524-525 (2023) 204838. doi:10.1016/j.wear.2023.204838.
[61] A. Kupin, Neural Identification of Technological Process of Iron Ore Beneficiation, in: 2007 4th IEEE Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, 2007, pp. 225–227. doi:10.1109/IDAACS.2007.4488409.
[62] A. Kupin, Research of properties of conditionality of task to optimization of processes of concentrating technology is on the basis of application of neural networks, Metallurgical and Mining Industry 6 (2014) 51–55.
[63] A. V. Morozov, T. A. Vakaliuk, I. A. Tolstoy, Y. O. Kubrak, M. G. Medvediev, Digitalization of thesis preparation life cycle: a case of zhytomyr polytechnic state university, CEUR Workshop Proceedings 3553 (2023) 142–154. URL: https://ceur-ws.org/Vol-3553/paper14.

pdf.
[64] R. P. Kukharchuk, T. A. Vakaliuk, O. V. Zaika, A. V. Riabko, M. Medvediev, Implementation of STEM learning technology in the process of calibrating an NTC thermistor and developing an electronic thermometer based on it, CEUR Workshop Proceedings 3358 (2022) 39–52. URL: https://ceur-ws.org/Vol-3358/paper25.pdf.
[65] N. Balyk, O. Barna, G. Shmyger, V. Oleksiuk, Model of Professional Retraining of Teachers Based on the Development of STEM Competencies, CEUR Workshop Proceedings 2104 (2018) 318–331. URL: https://ceur-ws.org/Vol-2104/paper_157.pdf.
[66] O. Spirin, V. Oleksiuk, O. Oleksiuk, S. Sydorenko, The Group Methodology of Using Cloud Technologies in the Training of Future Computer Science Teachers, CEUR Workshop Proceedings 2104 (2018) 294–304. URL: https://ceur-ws.org/Vol-2104/paper_154.pdf.
[67] S. Semerikov, S. Chukharev, S. Sakhno, A. Striuk, A. Iatsyshyn, S. Klimov, V. Osadchyi, T. Vakaliuk, P. Nechypurenko, O. Bondarenko, H. Danylchuk, Our sustainable pandemic future, E3S Web of Conferences 280 (2021) 00001. doi:10.1051/e3sconf/202128000001.
[68] D. S. Shepiliev, Y. O. Modlo, Y. V. Yechkalo, V. V. Tkachuk, M. M. Mintii, I. S. Mintii, O. M. Markova, T. V. Selivanova, O. M. Drashko, O. O. Kalinichenko, T. A. Vakaliuk, V. V. Osadchyi, S. O. Semerikov, Webar development tools: An overview, CEUR Workshop Proceedings 2832 (2020) 84–93.
[69] S. A. MacGowan, F. Madeira, T. Britto-Borges, M. Warowny, A. Drozdetskiy, J. B. Procter, G. J. Barton, The Dundee Resource for Sequence Analysis and Structure Prediction, Protein Science 29 (2020) 277–297. doi:10.1002/pro.3783.
[70] H. Wright, K. Brodlie, J. Wood, J. Procter, Problem Solving Environments: Extending the Rôle of Visualization Systems, in: A. Bode, T. Ludwig, W. Karl, R. Wismüller (Eds.), Euro-Par 2000 Parallel Processing, Springer Berlin Heidelberg, Berlin, Heidelberg, 2000, pp. 1323–1331. doi:10.1007/3-540-44520-X_185.
[71] V. Derbentsev, S. Semerikov, O. Serdyuk, V. Solovieva, V. Soloviev, Recurrence based entropies for sustainability indices, E3S Web of Conferences 166 (2020) 13031. doi:10.1051/e3sconf/202016613031.
[72] A. Kiv, V. Soloviev, S. Semerikov, H. Danylchuk, L. Kibalnyk, A. Matviychuk, Experimental economics and machine learning for prediction of emergent economy dynamics, CEUR Workshop Proceedings 2422 (2019) 1–4.
[73] A. Ganbayev, E. Seyidzade, Enhancing Customs Fraud Detection: A Comparative Study of Methods for Performance Measurement and Feature Improvement, in: 2023 IEEE 17th International Conference on Application of Information and Communication Technologies (AICT), 2023, pp. 1–5. doi:10.1109/AICT59525.2023.10313153.
[74] A. Adamov, S. Mehdiyev, E. Seyidzade, Good practice of data modeling and database design for UMIS. Course registration system implementation, in: 2014 IEEE 8th International Conference on Application of Information and Communication Technologies (AICT), 2014, pp. 1–4. doi:10.1109/ICAICT.2014.7035949.
[75] S. O. Semerikov, A. M. Striuk, T. A. Vakaliuk, A. Morozov, Quantum information technology on the Edge, CEUR Workshop Proceedings 2850 (2021) 1–15. URL: http://ceur-ws.org/Vol-2850/paper0.pdf.
[76] S. O. Semerikov, S. M. Chukharev, S. I. Sakhno, A. M. Striuk, A. V. Iatsyshin, S. V. Klimov,

V. V. Osadchyi, T. A. Vakaliuk, P. P. Nechypurenko, O. V. Bondarenko, H. B. Danylchuk, 3rd International Conference on Sustainable Futures: Environmental, Technological, Social and Economic Matters, IOP Conference Series: Earth and Environmental Science 1049 (2022) 011001. doi:10.1088/1755-1315/1049/1/011001.

[77] T. A. Vakaliuk, L. D. Shevchuk, B. V. Shevchuk, Possibilities of using AR and VR technologies in teaching mathematics to high school students, Universal Journal of Educational Research 8 (2020) 6280 – 6288. doi:10.13189/ujer.2020.082267.

[78] T. Vakaliuk, D. Antoniuk, A. Morozov, M. Medvedieva, M. Medvediev, Green IT as a tool for design cloud-oriented sustainable learning environment of a higher education institution, E3S Web of Conferences 166 (2020) 10013. doi:10.1051/e3sconf/202016610013.

[79] V. Voytenko, Some challenges in mobile context-aware applications for courses in academia, in: N. C. Callaos, B. Sanchez, H. W. Chu, J. Ferrer, S. L. Fernandes (Eds.), 7th International Multi-Conference on Complexity, Informatics and Cybernetics, IMCIC 2016 and 7th International Conference on Society and Information Technologies, ICSIT 2016 - Proceedings, volume 1, International Institute of Informatics and Systemics, IIIS, 2016, pp. 244–245.

[80] F. Lin, A. Dewan, V. Voytenko, Open Interactive Algorithm Visualization, in: 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE), 2019, pp. 1–4. doi:10.1109/CCECE.2019.8861535.

[81] O. V. Bondarenko, P. P. Nechypurenko, V. A. Hamaniuk, S. O. Semerikov, Educational Dimension: a new journal for research on education, learning and training, Educational Dimension 1 (2019) 1–4. doi:10.31812/ed.620.

[82] S. Semerikov, Educational Technology Quarterly: in the beginning, Educational Technology Quarterly 2021 (2021) 1–50. doi:10.55056/etq.13.

[83] S. Papadakis, A. E. Kiv, H. M. Kravtsov, V. V. Osadchyi, M. V. Marienko, O. P. Pinchuk, M. P. Shyshkina, O. M. Sokolyuk, I. S. Mintii, T. A. Vakaliuk, L. E. Azarova, L. S. Kolgatina, S. M. Amelina, N. P. Volkova, V. Y. Velychko, A. M. Striuk, S. O. Semerikov, ACNS Conference on Cloud and Immersive Technologies in Education: Report, CTE Workshop Proceedings 10 (2023) 1–44. doi:10.55056/cte.544.

[84] T. A. Vakaliuk, Editorial for JEC Volume 2 Issue 2 (2023), Journal of Edge Computing 2 (2023) 102–103. doi:10.55056/jec.654.

[85] T. A. Vakaliuk, S. O. Semerikov, Introduction to doors Workshops on Edge Computing (2021-2023), Journal of Edge Computing 2 (2023) 1–22. doi:10.55056/jec.618.

[86] A. I. Jony, A. K. B. Arnob, A long short-term memory based approach for detecting cyber attacks in IoT using CIC-IoT2023 dataset, Journal of Edge Computing (2024). doi:10.55056/jec.648.

[87] I. A. Pilkevych, D. L. Fedorchuk, M. P. Romanchuk, O. M. Naumchak, Approach to the fake news detection using the graph neural networks, Journal of Edge Computing 2 (2023) 24–36. doi:10.55056/jec.592.

[88] N. M. Lobanchykova, I. A. Pilkevych, O. Korchenko, Analysis and protection of IoT systems: Edge computing and decentralized decision-making, Journal of Edge Computing 1 (2022) 55–67. doi:10.55056/jec.573.

[89] N. Balyk, S. Leshchuk, D. Yatsenyak, Design and implementation of an IoT-based educational model for smart homes: a STEM approach, Journal of Edge Computing 2 (2023)

148–162. doi:10.55056/jec.632.

[90] A. V. Ryabko, O. V. Zaika, R. P. Kukharchuk, T. A. Vakaliuk, Graph theory methods for fog computing: A pseudo-random task graph model for evaluating mobile cloud, fog and edge computing systems, Journal of Edge Computing 1 (2022) 1–16. doi:10.55056/jec.569.

[91] T. A. Uzdenov, A new approach for dispatching task flows in GRID systems with inalienable resources, Journal of Edge Computing 1 (2022) 68–80. doi:10.55056/jec.574.

[92] A. V. Riabko, T. A. Vakaliuk, O. V. Zaika, R. P. Kukharchuk, V. V. Kontsedailo, Investigating the effect of virtual machine migration accounting on reliability using a cluster model, Journal of Edge Computing 2 (2023) 37–63. doi:10.55056/jec.585.

[93] O. V. Talaver, T. A. Vakaliuk, Reliable distributed systems: review of modern approaches, Journal of Edge Computing 2 (2023) 84–101. doi:10.55056/jec.586.

[94] T. Lorido-Botran, M. K. Bhatti, ImpalaE: Towards an optimal policy for efficient resource management at the edge, Journal of Edge Computing 1 (2022) 43–54. doi:10.55056/jec.572.

[95] M. V. Klymenko, A. M. Striuk, Design and implementation of an edge computing-based GPS tracking system, Journal of Edge Computing 2 (2023) 175–189. doi:10.55056/jec.634.

[96] A. R. Petrosian, R. V. Petrosyan, I. A. Pilkevych, M. S. Graf, Efficient model of PID controller of unmanned aerial vehicle, Journal of Edge Computing 2 (2023) 104–124. doi:10.55056/jec.593.

[97] T. M. Nikitchuk, T. A. Vakaliuk, O. A. Chernysh, O. L. Korenivska, L. A. Martseva, V. V. Osadchyi, Non-contact photoplethysmographic sensors for monitoring students' cardiovascular system functional state in an IoT system, Journal of Edge Computing 1 (2022) 17–28. doi:10.55056/jec.570.

[98] T. M. Nikitchuk, O. V. Andreiev, O. L. Korenivska, M. G. Medvediev, Model of an automated biotechnical system for analyzing pulseograms as a kind of edge devices, Journal of Edge Computing 2 (2023) 64–83. doi:10.55056/jec.627.

[99] O. L. Korenivska, V. B. Benedytskyi, O. V. Andreiev, M. G. Medvediev, A system for monitoring the microclimate parameters of premises based on the Internet of Things and edge devices, Journal of Edge Computing 2 (2023) 125–147. doi:10.55056/jec.614.

[100] A. G. Tkachuk, M. S. Hrynevych, T. A. Vakaliuk, O. A. Chernysh, M. G. Medvediev, Edge computing in environmental science: automated intelligent robotic platform for water quality assessment, Journal of Edge Computing 2 (2023) 163–174. doi:10.55056/jec.633.

An approach to assessing the reliability of software systems based on a graph model of method dependence

Valerii Krutko, Iryna Spivak and Svitlana Krepych

West Ukrainian National University, 11 Lvivska Str., Ternopil, 46009, Ukraine

Abstract

The paper deals with one of the main software development tasks, namely, the research of one of its main quality criteria - reliability. The analysis of recent studies of the subject has shown that architecture-based software reliability assessment methods are more informative and adequate. However, their accuracy in many cases will depend on the complexity of the software architecture. The authors of the paper propose an approach to assessing the reliability of software systems based on a graph model of method dependency. This approach includes elements of structural analysis of the source code of the system under study in order to build method dependency graph to assess its reliability based on stochastic reliability indicators of each method. The overall reliability calculation takes into account the complexity of the software system. Based on this approach, an intellectualized system for assessing the reliability of software systems with different architectures has been developed and tested on the example of two systems of different complexity.

Keywords

software system, reliability, evaluation, stochastic indicators, graph model, probability

1. Introduction

In today's world, software has become an integral part of many areas of our daily lives, from ensuring the efficient operation of businesses to creating comfort for individuals. The development market is also rapidly changing and adapting to meet the growing demand for new software. And the key aspects of these changes are, first of all, the desire to increase the speed of program development and reduce the final price of the software product.

It is well known that high quality software is an integral part of a successful product. However, even with a fairly slow development pace, developers often make mistakes that lead to serious problems, affecting security, reliability, and user satisfaction, not to mention development in a short time. That is why ensuring the high quality of a software product is one of the main tasks that must be solved at the development stage [1].

CS&SE@SW 2023: 6th Workshop for Young Scientists in Computer Science & Software Engineering, February 2, 2024, Kryvyi Rih, Ukraine

jave123@ukr.net (V. Krutko); spivak.iruna@gmail.com (I. Spivak); msya220189@gmail.com (S. Krepych)
https://www.wunu.edu.ua/educational-subdivisions/faculty/fkit/department-kn-fkit/staff-kn-fkit/7027-spivak-iryna-yaroslavivna.html (I. Spivak); https://www.wunu.edu.ua/educational-subdivisions/faculty/fkit/department-kn-fkit/staff-kn-fkit/8416-krepych-svitlana-yaroslavivna.html (S. Krepych)
0000-0003-4831-0780 (I. Spivak); 0000-0001-7700-8367 (S. Krepych)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

CEUR-WS.org/Vol-3662/paper11.pdf

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

From the point of view of reducing the cost of development and shortening its time, it is important to identify mistakes as early as possible since in iterative development, the complexity, and as a result the cost of mistake fixing, increases over time, and the probability of fixing it correctly (without introducing new mistakes) decreases [2]. On the other hand, almost all large software products are too complex to contain no mistakes at all. Considering mentioned above, the development task is not to guarantee the absence of mistakes, but to maintain a balance between the conditional reliability of the final software system and the costs at all development stages.

All existing software reliability assessment methods can be divided into three categories [3, 4]:

- black box reliability models;
- software metric based reliability models;
- architecture-based reliability models.

The first two categories are technically not reliability evaluation methods, but rather failure prediction methods, as they are based on the paradigms of increasing reliability and failure rate distribution over time. Since the probability of failure for software systems does not directly depend on the operating time [3], such an estimate is rather inaccurate. The third type of models is a more attractive alternative to the previous ones, as it clearly links program reliability to component reliability, which makes it easier to identify components that are critical in terms of reliability. In addition, these methods can be used for early program reliability assessment [5].

The analysis of available architecture-based reliability models has shown that they are either too complex to be used by ordinary programmers [5], or use entire software modules as architectural units [6, 7], which has a negative impact on accuracy.

The purpose of this paper is to develop an approach that would allow automated assessment of the reliability of software systems without the need to calculate and validate auxiliary coefficients. In addition, this approach is an attempt to move away from the rather traditional strategy of adapting hardware reliability methods to assess software systems.

2. Problem statement

To solve the problem of assessing the reliability of software systems, a number of reliability models are currently used [8]. However, all of these models are partially or fully based on models for assessing the reliability of hardware systems with minor changes, so they contain a number of simplifications and assumptions, which significantly limits the possibility of their application with real software products.

The difference between software and hardware systems is that in software systems, the probability of detecting an error does not directly depend on time, since if a static set of data is input, the program will work without fail indefinitely. This leads to the fact that expressing reliability in terms of time between failures evaluates the tester's qualifications rather than the actual reliability of the program.

Although mistakes are related to the concept of reliability, they cannot be a unit of measurement of reliability by themselves because, even when divided by the number of structural

elements of the program, they do not characterize the ability (or inability) of a software product to convert a set of input data into the expected result of execution. In addition, the number of mistakes in a program does not directly affect reliability, since:

- the result of observing the system is a failure, while the mistake remains hidden;
- failure can be caused by several mistakes simultaneously;
- mistakes can compensate each other in a way that after correcting one of them, the number of failures increases;
- some mistakes can be manifested only in the case of input data set that is impossible or extremely unlikely to be generated under real operating conditions;
- some mistakes lead to a false result regardless of the input data, so in fact, one mistake can cause zero program reliability.

3. Description of the proposed method

A flowchart of the approach to assessing the reliability of software systems based on a graph model of method dependence is shown in figure 1.

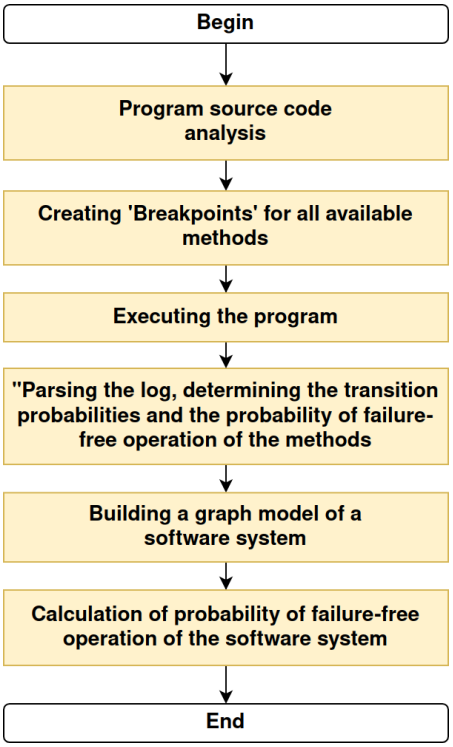


Figure 1: Flowchart of the approach to assessing the reliability of software systems based on a graph model of method dependence.

So, the first step of this approach is “Program source code analysis”, the purpose of which is to break down the program system as a whole into smaller structural elements – methods. To

evaluate the reliability of such a system, we use a graph model [9] in which the vertices of the graph are methods, and the edges are the links between these methods in the program under study. Figure 2 shows an example of such a graph model for a software system consisting of four methods.

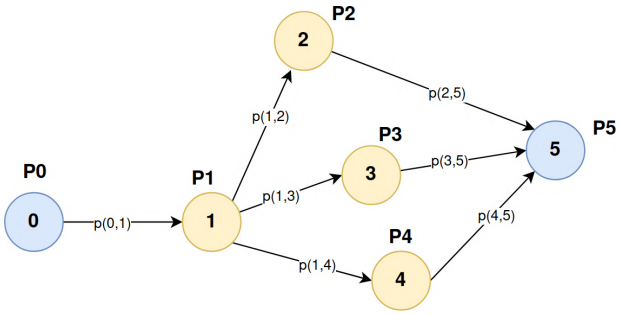


Figure 2: Graph model of the program consisting of four methods.

This oriented graph contains 6 vertices, four of which are responsible for methods, and another two are dummy vertices representing points of entry to and points of exit from the program code.

The values P_1, P_2, P_3 and P_4 are the probabilities of failure-free operation of the corresponding methods, and $p(i, j)$ are the probabilities of invocation of the j-th method from the body of the i-th method.

The probability of failure-free operation of each method is calculated as the ratio of the number of its invocations that caused an exception throwing to the total number of method invocations. The transition probabilities are calculated in proportion to the number of invocations to nested methods from the body of the i-th method, so that the sum of all transition probabilities from the body of the i-th method is one (1):

$$\sum_{j=j_{min}}^{j_{max}} p(i, j) = 1 \tag{1}$$

These probabilities are calculated by performing the following steps of the proposed approach, namely “Creating ‘Breakpoints’ for all available methods”, “Executing the program” and “Parsing the log, determining the transition probabilities and the probability of failure-free operation of the methods”.

The next step of the approach is “Building a graph model of a software system”. First of all, a transition probability matrix P_1 with dimension $m \times m$ is constructed, where m is the number of vertices in the graph. In this matrix, the value of the element $P(i, j)$ is the value of the probability of the corresponding transition between the corresponding methods, i.e. $p(i, j)$. If there is no relationship between the corresponding methods, then zero is written in place of the corresponding element. The next step is to calculate the matrix G, for which the j-th row of the matrix P is multiplied by the probability of failure-free operation of the j-th method (for dummy vertices, this parameter is equal to one).

Finally, it is necessary to calculate the matrix T, which in the case of an acyclic graph is calculated using formula (2).

$$T = I + G + G^2 + \cdots + G^m \tag{2}$$

And in the case of non-acyclic, by formula (3).

$$T = (I - G)^{-1} \tag{3}$$

The probability of failure-free operation of the software system under test, which is calculated at the last step of the approach, is the value of the element $T_{0,m-1}$ of the matrix T.

4. An example of application of the proposed approach for assessing the reliability of a software system based on a graph model of method dependence

To illustrate the implementation of the proposed approach for assessing the reliability of a software system based on a graphical model of method dependency, the reliability assessment of a simple program has been illustrated, the program code of which is shown in figure 3.

```
public class Main {
    private static int i = 0;

    public static void main(String[] args){
        for(; i++ < Integer.MAX_VALUE;){
            try{
                calculateSomething();
            } catch (Exception ignored){}
        }
    }

    public static void calculateSomething(){
        double a = generateDividend() / generateDivisor();
    }

    public static double generateDividend(){
        if(i % 21 == 0)
            throw new NumberFormatException("Something happened");
        if(i % 3 == 0) return genRandomValue();
        else return getConstantValue();
    }

    private static double genRandomValue(){
        if(i % 99 == 0)
            throw new NumberFormatException("Something happened");
        return (Math.random() * (Math.random() > 0.5d ? 1 : -1));
    }

    private static double getConstantValue(){
        if(i % 95 == 0)
            throw new NumberFormatException("Something happened");
        return 1.751;
    }

    public static Double generateDivisor(){
        if(i % 75 == 0) return null;
        return Math.random();
    }
}
```

Figure 3: Source code of the program under test.

For the convenience of visual demonstration, all methods of a software system are located in the same class, but the proposed approach will work even if all methods belong to different classes.

To emulate the presence of mistakes, conditional statements have been added to the bodies of methods that throw exceptions of the `NumberFormatException` type at certain values of the iterative variable i. The linkage to the iterative variable was made solely in terms of repeatability of results for ease of testing. Building a graph model is started here by creating a methods invocation tree. The tree structure of the program under study in JSON format is shown in figure 4.


```
{
  "name": "calculateSomething",
  "location": "Main.java",
  "nestedMethods": [{
      "name": "generateDividend",
      "location": "Main.java",
      "nestedMethods": [{
          "name": "genRandomValue",
          "location": "Main.java",
          "nestedMethods": []
        },{
          "name": "getConstantValue",
          "location": "Main.java",
          "nestedMethods": []
        }]
    },{
      "name": "generateDivisor",
      "location": "Main.java",
      "nestedMethods": []
    }]
}
```


Figure 4: Structure of the program under study in JSON format.

As figure 4 shows, the program's structural model contains the "location" field, meaning that this approach will work correctly with methods of the same name located in different classes.

To estimate the number of method executions and the number of failures, Intellij Idea's debug mode is applied to make it log all entries into the methods and all exceptions thrown [10]. 'Breakpoints' settings are shown in figure 5.

The log contains all the information necessary for further analysis. After processing this data, we can calculate the probability of failure for each method. The results of the calculations are shown in table 1.

Using the information from table 1, we can build a graph model of the system under study, as illustrated in figure 6.

As it can be seen, the graph shown in figure 6 does not contain cycles, so calculations will be performed by using formula (2). The graph contains seven vertices, so the matrices P, G and T will be of dimension 7×7. Figure 7 shows the values of all matrices.

As noted above, the probability of system's failure-free operation is determined by the element $T_{0,m-1}$ of the matrix T, since in our case $m = 7$, then $T_{0,m-1} = T_{0,6} = 0.93$

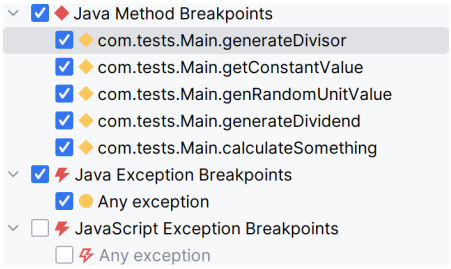


Figure 5: 'Breakpoints' settings.

Table 1
Log analysis results.

Method	Number of invocations	Number of failures	Probability of failure-free operation
calculateSomething()	2 147 483 647	23 798 952	0.9889
generateDividend()	2 147 483 647	102 261 126	0.9524
genRandomValue()	613 566 756	18 592 932	0.9697
getConstantValue()	1 431 655 765	15 070 061	0.9895
generateDivisor()	2 011 559 528	0	1

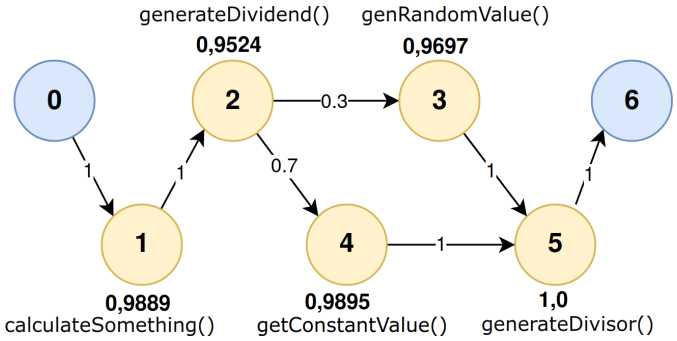


Figure 6: Graph model of a software system.

P

0.0	1.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	1.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.3	0.7	0.0	0.0
0.0	0.0	0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	1.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0

G

0.0	1.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.99	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.29	0.67	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.97	0.0
0.0	0.0	0.0	0.0	0.0	0.99	0.0
0.0	0.0	0.0	0.0	0.0	0.0	1.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0

T

1.0	1.0	0.99	0.28	0.66	0.93	0.93
0.0	1.0	0.99	0.28	0.66	0.93	0.93
0.0	0.0	1.0	0.29	0.67	0.94	0.94
0.0	0.0	0.0	1.0	0.0	0.97	0.97
0.0	0.0	0.0	0.0	1.0	0.99	0.99
0.0	0.0	0.0	0.0	0.0	1.0	1.0
0.0	0.0	0.0	0.0	0.0	0.0	1.0

Figure 7: The values of the matrices P, G and T.

Test program whose code is shown in the figure 1 is simple and contains only one entry point. It was made intentionally to be able to check calculations results using traditional methods. According to the fundamentals of reliability theory [11], the probability of system failure is defined as:

$$P(t) = 1 - \frac{n(t)}{N_0}$$

where N_0 is the number of elements at the beginning of the test and $n(t)$ the number of elements that failed during the time interval t.

As it can be seen, program invokes the `calculateSomething()` method 2 147 483 647 times (`Integer.MAX_VALUE`). The number of application failures in this case would be equal to the total number of generated exceptions. Using the data from Table 1, system's failure-free operation probability can be calculated as:

$$P = 1 - \frac{23\,798\,952 + 102\,261\,126 + 18\,592\,932 + 15\,070\,061 + 0}{2\,147\,483\,647} = \\ = \frac{159\,723\,071}{2\,147\,483\,647} = 0.9256$$

As it can be seen, the results match, which indicates that proposed approach works correctly.

5. An example of using the method for more complex systems

Now let's apply the proposed method to measure the reliability of a more complex program. The class diagram of the test program is shown in the figure 8.

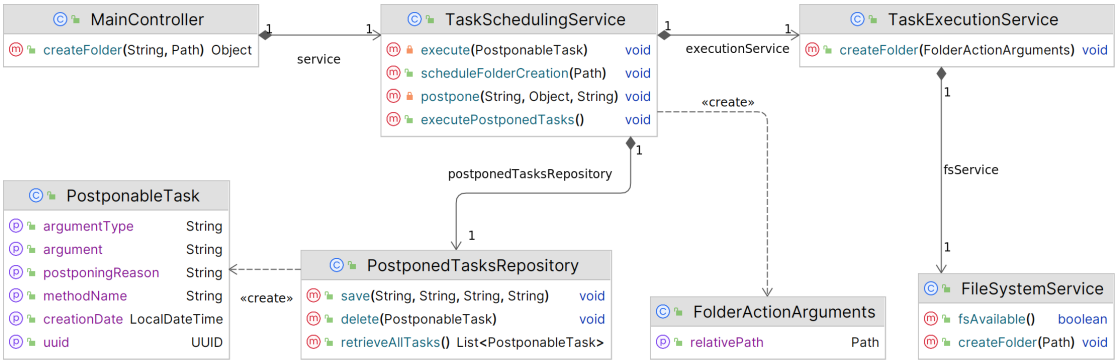


Figure 8: Test application class diagram.

This program is a web application that was created using the Spring framework. The program contains 7 classes and 11 methods and 2 entry points – the first is implemented through the controller, and the second through the task scheduler.

As in the previous example, we will get information about the number of method invocations and errors that have occurred by parsing the log. Structured data are shown in table 2.

As it can be seen, there are three methods `createFolder()` in the source code of the program, but they were separated by the model and processed correctly. Figure 9 shows a graphical model of the system under study, built using data from table 2.

Table 2
Method invocation statistics.

Method	Number of invocations	Number of failures	Probability of failure-free operation
MainController.createFolder()	1 000	0	1
scheduleFolderCreation()	1 000	0	1
TaskExecutionService.createFolder()	3765	0	1
fsAvailable()	3765	35	0.991
postpone()	333	5	0.985
save()	328	2	0.993
FileSystemService.createFolder()	1074	75	0.93
delete()	326	1	0.997
executePostponedTasks()	114	0	1
retriveAllTasks()	114	2	0.982
execute()	2768	13	0.995

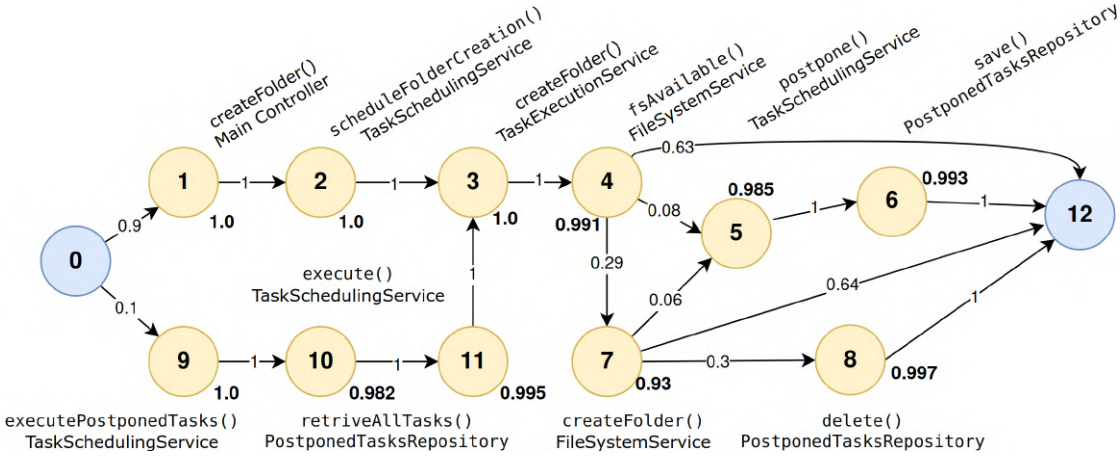


Figure 9: Graph model of a software system under study.

Since graph contains 13 vertices, matrices P, G and T will be of dimension 13×13. Figure 10 shows the values of all matrices.

As in the previous example, the probability of system's failure-free operation is determined by the element $T_{0,m-1}$ and therefore $T_{0,m-1} = T_{0,12} = 0.97$.

From the perspective of model performance, parsing log files turned out to be the most difficult task, as expected. Since the sizes of the logs in the first and second examples differ significantly, it would be more appropriate to compare the time spent on calculations. The time to calculate the matrices for the first and second examples was 2.2ms and 3.6ms, respectively. In general, the calculation time (including log parsing) did not exceed 50ms in both cases. Taking this into account, it can be concluded that the proposed model can calculate reliability indicators for much larger software systems in an acceptable amount of time.

P

0.0	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0
0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.08	0.0	0.29	0.0	0.0	0.0	0.0	0.63
0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
0.0	0.0	0.0	0.0	0.0	0.06	0.0	0.0	0.3	0.0	0.0	0.0	0.64
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

G

0.0	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0
0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.08	0.0	0.29	0.0	0.0	0.0	0.0	0.62
0.0	0.0	0.0	0.0	0.0	0.0	0.99	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.99
0.0	0.0	0.0	0.0	0.0	0.06	0.0	0.0	0.28	0.0	0.0	0.0	0.6
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.98	0.0
0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

T

1.0	0.9	0.9	1.0	1.0	0.1	0.09	0.29	0.08	0.1	0.1	0.1	0.97
0.0	1.0	1.0	1.0	1.0	0.1	0.09	0.29	0.08	0.0	0.0	0.0	0.97
0.0	0.0	1.0	1.0	1.0	0.1	0.09	0.29	0.08	0.0	0.0	0.0	0.97
0.0	0.0	0.0	1.0	1.0	0.1	0.09	0.29	0.08	0.0	0.0	0.0	0.97
0.0	0.0	0.0	0.0	1.0	0.1	0.09	0.29	0.08	0.0	0.0	0.0	0.97
0.0	0.0	0.0	0.0	0.0	1.0	0.99	0.0	0.0	0.0	0.0	0.0	0.98
0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.99
0.0	0.0	0.0	0.0	0.0	0.06	0.05	1.0	0.28	0.0	0.0	0.0	0.93
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	1.0
0.0	0.0	0.0	0.98	0.98	0.09	0.09	0.28	0.08	1.0	1.0	0.98	0.95
0.0	0.0	0.0	0.98	0.98	0.09	0.09	0.28	0.08	0.0	1.0	0.98	0.95
0.0	0.0	0.0	1.0	1.0	0.09	0.09	0.29	0.08	0.0	0.0	1.0	0.96
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0

Figure 10: The values of the matrices P, G and T.

6. Conclusions

To test the proposed approach, it was implemented as an plug-in for integrated development environment Intellij Idea which could measure the reliability of programs written in Java. With minor changes, it can be adapted to work with code written in any object-oriented programming language. The objective was to develop an approach that would make it possible to evaluate the reliability of both – the entire software system and its specific modules at the testing stage. Upon testing completion, reliability indicator is calculated, which could be used to make decisions about the feasibility of further program improvement.

The example shows the effectiveness of this approach. The developed intellectualized system based on the proposed approach can be used to assess software systems reliability of any complexity and architecture.

References

[1] ISO/IEC 25010:2011, Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models, Standard, International Organization for Standardization, 2011.

[2] V. V. Vyshnivskyi, V. V. Vasylenko, M. P. Hnidenko, Osnovy nadiinosti ta diahnostyky informatsiinykh system [Fundamentals of reliability and diagnostics of information systems], FOP Huliaieva V. M., 2020.

[3] I. Eusgeld, F. Fraikin, M. Rohr, F. Salfner, U. Wappler, Software Reliability, in: I. Eusgeld, F. C. Freiling, R. Reussner (Eds.), Dependability Metrics: Advanced Lectures, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008, pp. 104–125. doi:10.1007/978-3-540-68947-8_10.

[4] N. R. Barraza, Five Decades of Software Reliability, Past, Present, Future and New Challenges, in: 2018 7th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2018, pp. 88–94. doi:10.1109/ICRITO.2018.8748556.

[5] A. Dimov, S. Punnekkat, Fuzzy Reliability Model for Component-Based Software Systems, in: 2010 36th EUROMICRO Conference on Software Engineering and Advanced Applications, 2010, pp. 39–46. doi:10.1109/SEAA.2010.46.

[6] K. Goševa-Popstojanova, K. S. Trivedi, Architecture-based approach to reliability assessment of software systems, Performance Evaluation 45 (2001) 179–204. doi:https://doi.org/10.1016/S0166-5316(01)00034-7, Performance Validation of Software Systems.

[7] V. C. Skanda, S. Srinivasa Prasad, G. R. Dheemanth, N. S. Kumar, Assessment of Quality of Program Based on Static Analysis, in: 2019 IEEE Tenth International Conference on Technology for Education (T4E), 2019, pp. 276–277. doi:10.1109/T4E.2019.00072.

[8] V. Yakovyna, V. Matseliukh, Ohliad i analiz modelei nadiinosti prohramnoho zabezpechennia [Review and analysis of software reliability models], Visnyk Natsionalnoho universytetu "Lvivska politekhnika". Kompiuterni nauky ta informatsiini tekhnolohii 864 (2017) 103–138. URL: https://science.lpnu.ua/sites/default/files/journal-paper/2018/jul/13773/17.pdf.

[9] S. Krepych, I. Spivak, S. Spivak, Methodology of Formation of the Individual Study Plan of the Student Based on the Graph Model of the Dependence of Disciplines, CEUR Workshop Proceedings 3426 (2023) 298–307. URL: https://ceur-ws.org/Vol-3426/paper24.pdf.

[10] Breakpoints | IntelliJ IDEA Documentation, 2024. URL: https://www.jetbrains.com/help/idea/using-breakpoints.html.

[11] Y. Yan, Y. Peng, C. Liu, H. Li, A New Software Reliability Evaluation Model Based on the Probability of the Failure Mode, in: 2018 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C), 2018, pp. 35–41. doi:10.1109/QRS-C.2018.00020.

Methodology of implementation of modern information systems at commercial enterprises

Yurii O. Chernukha[1], Oksana V. Klochko[1] and Tetiana P. Zuziak[1]

[1]Vinnytsia Mykhailo Kotsiubynskyi State Pedagogical University, 32 Ostrozhskogo Str., Vinnytsia, 21100, Ukraine

Abstract

The article discusses the features of implementing ERP (Enterprise Resource Planning) based on the AIM (Application Implementation Method) methodology. The process of implementing an IS (Information System) is described, taking into account many years of practical experience and developments in the countries of Eastern Europe. Features of individual stages of ERP implementation, problems of their implementation and recommendations on possibilities of solving the problem are considered. A survey of experts from various technical fields was conducted on the main problems of IS implementation, their conclusions were analyzed, taken into account and agreed upon. An experiment was carried out to analyze complex solutions using the T. L. Saaty method, consistency indices have been calculated. Special attention is paid to the preparation of project documents, the study and understanding of which is necessary for Professional Junior Bachelors, since these employees will be involved in the implementation of the ERP system at the enterprise on the part of the customer. Their main responsibilities will be directly related to the description of existing business processes and the results of implementation depend on the quality of their work.

Keywords

Application Implementation Method, Enterprise Resource Planning, information system, information system implementation methodologies, software maintenance, software requirements, business process modeling, learning technologies, Professional Junior Bachelor

1. Introduction

The development of ERP (Enterprise Resource Planning) systems implementation methodology, their study, and analysis is an important and relevant direction in the field of enterprise management. The methodology, first of all, depends on the complexity of the implementation of the ERP system, which can radically change the organization, including the processes, structure, technologies and culture of the company. Methodology studies help to understand the complexity of these changes and provide recommendations for their successful implementation, mitigate resistance, and ensure successful adaptation of personnel.

Secondly, research allows us to develop new or optimize existing methodologies to ensure maximum efficiency in using the ERP system, taking into account the specifics of the industry and business. Methodologies can help the adoption and use of new technologies such as

CS&SE@SW 2023: 6th Workshop for Young Scientists in Computer Science & Software Engineering, February 2, 2024, Kryvyi Rih, Ukraine

yurvin@gmail.com (Y. O. Chernukha); klochkoob@gmail.com (O. V. Klochko); zuzyak@ukr.net (T. P. Zuziak)

0009-0007-7465-118X (Y. O. Chernukha); 0000-0002-6505-9455 (O. V. Klochko); 0000-0002-5437-0272 (T. P. Zuziak)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

CEUR-WS.org/Vol-3662/paper12.pdf

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

artificial intelligence, data analytics, cloud computing, etc. Research helps determine how these innovations can be integrated into the enterprise management system.

Thirdly, due to the use of methodologies, risks such as budget overruns, delays in deadlines, data loss, etc. are significantly reduced.

Methodology research helps identify such potential risks and propose strategies to reduce them. Only the effective implementation of ERP systems can increase the competitiveness of an enterprise, providing better resource management, increased productivity and improved quality of decision-making. Therefore, the study of the methodology for implementing ERP systems remains relevant in order to cope with the challenges associated with enterprise management in today's dynamic business environment.

An ERP system is software that integrates and automates various business processes in an organization and helps manage company resources.

Researchers are actively investigating various implementations of ERP systems, both in commercial organizations and integrator companies. They not only draw on the conceptualization of multiple theories based on technology acceptance and innovation diffusion models, but also consider the policy implications of fully implementing and using ERP to achieve company goals [1]. Separate works are devoted to identifying factors that predict the adoption and implementation of ERP systems in developing countries [2]. Much attention is paid to creating for effective, practical IS implementation plans. Comprehensive principles are developed using a conceptual research method based on the study of literature on the integration of various concepts, project management at different organizational levels and strategic directions [3].

Some authors go deeper into analyzing the errors of individual projects. For example, Menon [4] examined the critical problems of implementing an ERP system based on the results of a qualitative study of one specific project with an analysis of all project documents. Ivanović and Marić [5] described the concept and content of digital business transformation and the impact of modern digital technologies on IS.

The decision about the implementation of information systems can be made on the basis of mathematical techniques, which represent a convenient and universal basis for modeling multi-criteria assessment of alternatives and deducing final priorities [6, 7]. Researchers closely monitor the volume and efficiency of customization of ERP systems [8] in order to avoid exceeding the budget and deadlines for developing new and tuning existing IS functionality. A lot of attention is paid for post-project work, exchange of knowledge and experience, which plays a central role in the development of enterprise software [9]. Specific sociotechnical measures that lead to sustainable and successful operation of an ERP system throughout its life cycle are discussed.

From a technical point of view, there are also studies focusing on architectural issues, such as the description of a three-dimensional design to evaluate the integration modules of the IS obtained from a sample of 68 ERP module implementations [10]. The relationship between the integration of each ERP module and the efficiency of business processes is assessed.

Based on the analysis of factors and causes of failures of IS integration, continuous work is being carried out to create new implementation models and expand existing ones [11]. But it should be noted that in the vast majority of the cases we are talking about complementing the existing methodology and using best practices in the field of project management and software engineering.

The purpose. This article describes the general plan for implementing an ERP system in Ukrainian realities based on the universal Oracle AIM methodology. Recommendations for using the software are given. The implementation features of each stage of the project are analyzed based on practical experience.

2. Selection of methods and diagnostics

As a rule, each ERP system developer company provides a methodology for implementing its system. Let's consider the methodology for implementing Oracle AIM products, which has been used for more than 20 years by information system integrators, and not only Oracle. AIM (Application Implementation Method) is actually a set of interconnected document templates that defines the business needs of an enterprise at the beginning of a project and manages them throughout the implementation process, recording all actions in documents. AIM divides the project into six phases [12] (figure 1):

1. Definition – determination of project resources, budget, charter, team, business needs, selection of system and architecture.
2. Operation Analysis – comparison of standard functionality and business requirements, identifying gaps and finding solutions to them.
3. Solution Design – development of detailed design to cover business requirements.
4. Build – performing system customizations, testing, data conversion, cross-system testing.
5. Transition – deployment of the system in the enterprise, training and testing by users.
6. Production – implementation of an information system and support.

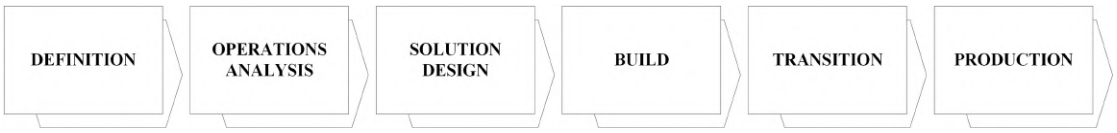


Figure 1: Six Project Phases.

There are 11 processes within the phases, which consist of documents [12]. Below is a list of documents for each of the 11 processes [12]:

1. Business process architecture (BP) (figure 2) – at this stage, a description of existing business practices, a catalog of changes, advanced and future practices is carried out.
2. Business requirements Definition (RD) (figure 3) – initial basic survey and collection of requirements, identification of business needs that should be covered by the implementation project.
3. Business Requirement Mapping (BR) (figure 4) – mapping future business requirements and standard information system capabilities to identify gaps.
4. Application and Technical Architecture (TA) (figure 5) – description of infrastructure requirements for the implementation of an information system, reflection of the business vision.

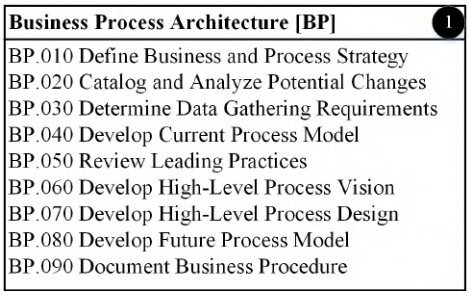


Figure 2: Business process architecture.

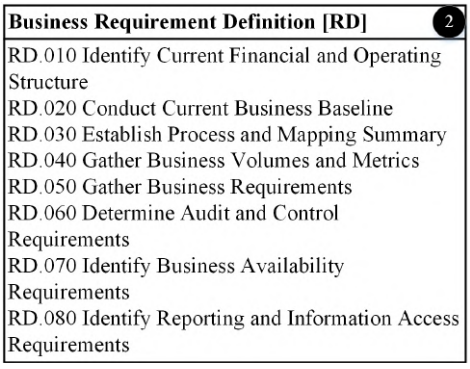


Figure 3: Business requirements Definition.

5. Module Design and Build (MD) (figure 6) – development of new customizations, describes in detail the design of the necessary forms, reports, alerts, database triggers, etc., which must be designed, created and tested before moving to a new system.
6. Data Conversions (CV) (figure 7) – data conversion, transformation or transfer of data from a previous system to a new information system.
7. Documentation (DO) (figure 8) – documentations for each module, including user manuals and implementation guides.
8. Business System Testing (TE) (figure 9) – the process of testing settings and customisations and standard functionality of the new system.
9. Performance Testing (PT) (figure 10) – stress testing (assessing transaction persistence time, transaction search time, workflow background processes, database performance, etc.).
10. Adoption and Learning (AP) (figure 11) – description of the reasons for the transition to a new system throughout the enterprise and assessment of the effectiveness of this transition.
11. Production Migration (PM) (figure 12) – a description of the decommissioning of the previous system and the implementation of the new one.

In order to evaluate the methodology of implementation of Information Systems at Commercial Enterprises, an expert evaluation was conducted.

The experts answered the following questionnaire questions:

1. Should the introduction of ERP into an enterprise take place with the participation of an integrator company?
2. Should the company dedicate trade secrets to the integrator and reflect such information in business process maps?
3. Can commercial enterprises fully adopt cloud technology?
4. Should the implementation of ERP in an enterprise (launch of all modules into commercial operation) last no more than three years?
5. Can an enterprise use ERP that is more than 10 years old?

Business Requirement Mapping [BR]	3
BR.010 Analyze High-Level Gaps BR.020 Prepare mapping environment BR.030 Map Business requirements BR.040 Map Business Data BR.050 Conduct Integration Fit Analysis BR.060 Create Information Model BR.070 Create Reporting Fit Analysis BR.080 Test Business Solutions BR.090 Confirm Integrated Business Solutions BR.100 Define Applications Setup BR.110 Define security Profiles	

Figure 4: Business Requirement Mapping.

Application and Technical Architecture [TA]	4
TA.010 Define Architecture Requirements and Strategy TA.020 Identify Current Technical Architecture TA.030 Develop Preliminary Conceptual Architecture TA.040 Define Application Architecture TA.050 Define System Availability Strategy TA.060 Define Reporting and Information Access Strategy TA.070 Revise Conceptual Architecture TA.080 Define Application Security Architecture TA.090 Define Application and Database Server Architecture TA.100 Define and Propose Architecture Subsystems TA.110 Define System Capacity Plan TA.120 Define Platform and Network Architecture TA.130 Define Application Deployment Plan TA.140 Assess Performance Risks TA.150 Define System Management Procedures	

Figure 5: Application and Technical Architecture.

Build and Module Design [MD]	5
MD.010 Define Application Extension Strategy MD.020 Define and estimate application extensions MD.030 Define design standards MD.040 Define Build Standards MD.050 Create Application extensions functional design MD.060 Design Database extensions MD.070 Create Application extensions technical design MD.080 Review functional and Technical designs MD.090 Prepare Development environment MD.100 Create Database extensions MD.110 Create Application extension modules MD.120 Create Installation routines	

Figure 6: Module Design and Build.

Data Conversion [CV]	6
CV.010 Define data conversion requirements and strategy CV.020 Define Conversion standards CV.030 Prepare conversion environment CV.040 Perform conversion data mapping CV.050 Define manual conversion procedures CV.060 Design conversion programs CV.070 Prepare conversion test plans CV.080 Develop conversion programs CV.090 Perform conversion unit tests CV.100 Perform conversion business objects CV.110 Perform conversion validation tests CV.120 Install conversion programs CV.130 Convert and verify data	

Figure 7: Data Conversions.

6. Should company employees who participate in the implementation of IS be allocated fully to work on the project?
7. Should the company's top management be directly involved in the project? Participate in meetings and make decisions on open issues?
8. Is it possible to provide on-the-job training to employees on the new IS?
9. Whether enterprise employees who do not belong to the project team can take part in testing the new IS?
10. Can artificial intelligence be involved in the implementation of IS?
11. Can modern IS support the adoption of machine learning methods and models?
12. Is there a risk of lack of professional growth of the company's employees and even degradation in case the implemented IS will take over the majority of current operations

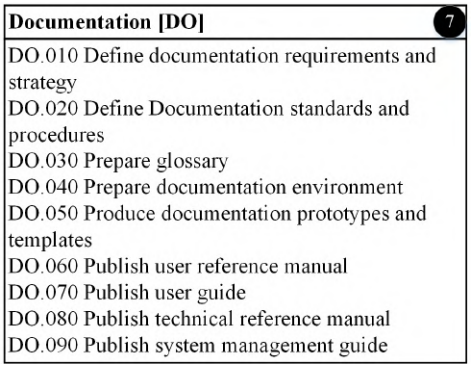


Figure 8: Documentation.

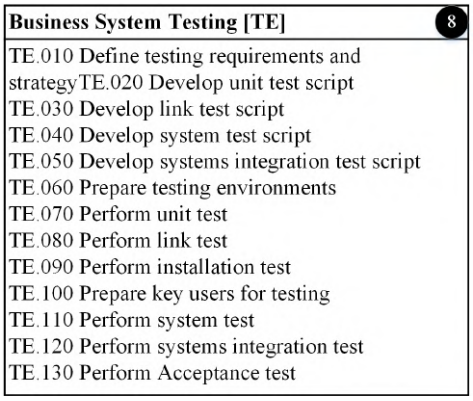


Figure 9: Business System Testing.

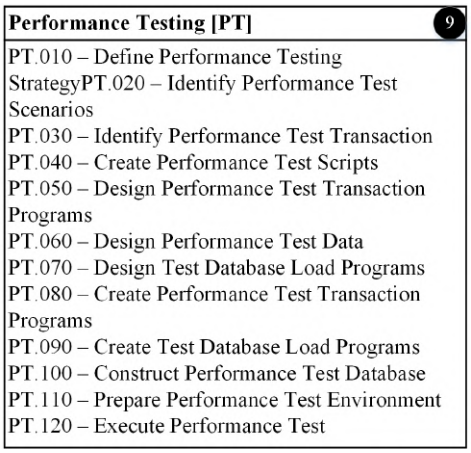


Figure 10: Performance Testing.

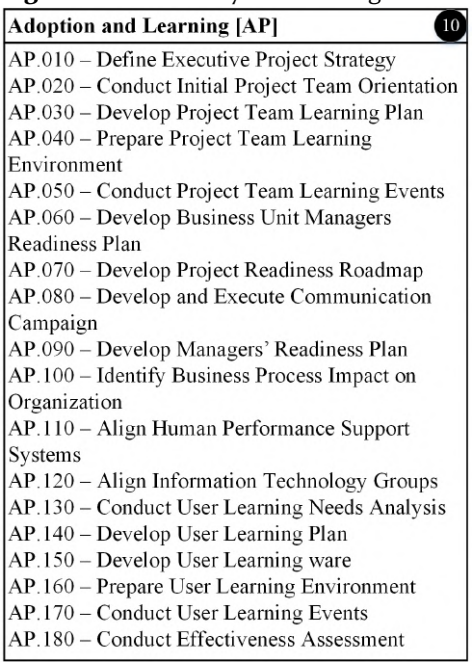


Figure 11: Adoption and Learning.

and user activities?

Answers to questionnaire questions are given on a scale from 1 to 10 points, where 1 is the lowest value and 10 is the highest value.

Consistency of experts' opinions was determined using the T. L. Saaty method [6]. Consistency Index (CI) was calculated according to formula (1) [6]. Consistency Ratio (CR) was calculated according to formula (2) [6].

$$CI = \frac{\lambda_{max} - n}{n - 1}, \tag{1}$$

where λ_{max} is the maximum eigenvalue of the matrix, n is the number of compared elements.

Production Migration [PM]	11
PM.010 – Define Transition Strategy PM.020 – Design Production Support Infrastructure PM.030 – Develop Transition and Contingency Plan PM.040 – Prepare Production Environment PM.050 – Set Up Applications PM.060 – Implement Production Support Infrastructure PM.070 – Verify Production Readiness PM.080 – Begin Production PM.090 – Measure System Performance PM.100 – Maintain System PM.110 – Refine Production System PM.120 – Decommission Former Systems PM.130 – Propose Future Business Direction PM.140 – Propose Future Technical Direction	

Figure 12: Production Migration.

$$CR = \frac{CI}{RI}, \tag{2}$$

where *RI* (Random Index) determined according to the Random Index table [6].

3. Results and discussion

The implementation of a new ERP system at an enterprise is a long-term, scrupulous process, the success of which is determined by the correct choice of the ERP system and the organization of the implementation process itself by project management. It requires careful planning, cooperation of all stakeholders and constant monitoring.

We will describe in more detail general stages of ERP system implementation:

1. Preparation and planning.
   1.1. Determining goals and needs, it is necessary to clearly understand what tasks and problems the ERP system should solve, what processes to cover, how the new system will be better than the previous one, what architecture the new ERP system will have.
   1.2. Creating a project team is an extremely responsible process. When creating a project team, it is necessary to carefully and individually approach the selection of technical and business architects, administrators, business analysts and developers. It is important to think through future communication with business process owners.
   1.3. Development of a project plan, a competent detailed implementation plan, determination of resources, budget and deadlines are the key to the future success of the project.
2. Selecting an ERP system.

2.1. Market analysis. It is necessary to study the global ERP systems market, find out the capabilities of each system, its architecture, support and updates. Understand how the best global practices that ERP systems provide can help an enterprise increase efficiency and solve existing problems.
2.2. The choice of an ERP system supplier is determined by the package of services provided, its experience and name in the market.
2.3. Selecting a consulting company that will be involved in implementation and/or consulting. This point can significantly speed up implementation, help with design and bring new practices to the company. The integrator company may provide support for the ERP system in the future.

3. Design.
3.1. Determining the general architecture of the ERP system, modules used, internal and external integrations and interfaces, creating an implementation plan.
3.2. Configuration is the setting of an ERP system in accordance with the needs of the company's business processes.
3.3. Creating requirements is documenting improvements or changes to the standard functionality of an ERP system to support current company practices that cannot be changed. This could also include changes to the user interface, integration with other applications.
3.4. Creating a project library that should be easy to use and update. All project documentation should be structured and stored in one place.

4. Development and testing.
4.1. Programming and settings of the ERP system, in accordance with paragraphs 3.2, 3.3.
4.2. Testing of completed settings and improvements, including functional, integration, stress and user testing.

5. Training and support.
5.1. Personnel training and preparation of user documentation.
5.2. Starting the system. Determined by the implementation option – "big bang" or phased implementation. Obviously, the second option is safer and more reliable for complex multi-module systems.

6. Analysis and optimization.
6.1. Monitoring, analysis and evaluation of the ERP system are daily mandatory work, the key to the reliability of the system's functioning.
6.2. Optimization and regular analyzes can significantly improve system performance and efficiency.

7. Support and update.
7.1. Technical support should be 24/7, and user support should cover all business hours of the enterprise.
7.2. Updates. Technical specialists and business analysts must monitor ERP system updates and support ongoing changes in the company's business processes.

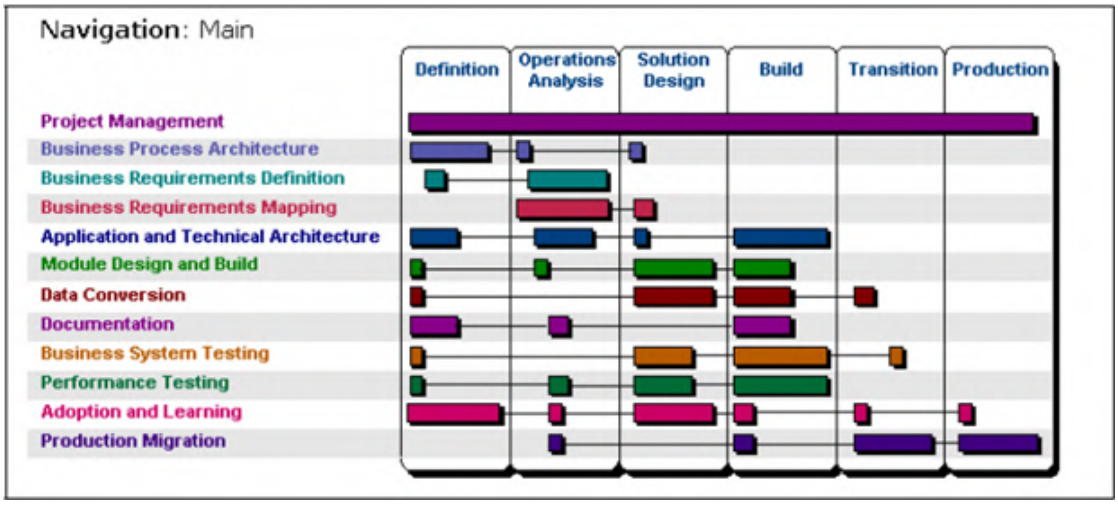


Figure 13: The time costs of implementing a new ERP system.

Time costs of implementing a new ERP system are displayed approximately as shown in the figure 13.

Using a list of documents for each of the 11 processes, you can use it to understand the essence of the documents (see figures 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12). Usually the necessary set is selected at the beginning of the project. There is no need to use the entire library. It is important to select significant documents for the current project, taking into account its complexity, timing, volume, etc. There is no need to create documents for the sake of filling the project library. The entire project team must clearly understand the list of project documents, documents purpose, and who is responsible for their creation.

There are no uniform standards for project activities when implementing ERP systems in Ukraine. Different software are used. Typically, project management determines the software used, guided by pragmatism, the complexity of the project and the level of the project team.

For communications, as a rule, an application is selected that has historically taken root in the enterprise. Because almost all departments of the company will be involved in the project to one degree or another. This could be Skype, Teams, Viber, Meet and others. It is desirable that the communicator program supports conferences and screen sharing.

The most popular tool for project management is Microsoft Project. It allows you to create detailed project plans, including tasks, dependencies, timelines, resources and budget. It is possible to use Gantt graphical tools to visualize the schedule of tasks and the dependencies between them. The program provides tools for assigning and managing project resources such as employees, equipment, and materials. It is convenient for a project manager to track the completion of all project tasks in Microsoft Project. The program also supports the ability to create reports and analytical information about the project. There is the possibility of integration and exchanges with other products.

Jira is often used for comfortable interaction between project team members. It is a popular project management system developed by Atlassian. Jira was originally created to track and

manage tasks in the software development industry. But then Jira has become used in various fields for project and task management.

The project library often chosen is Confluence, a modern collaboration and content management platform also created by Atlassian. Confluence is designed to facilitate shared work within teams by providing tools for creating, organizing and discussing different types of content. Confluence is a wiki-like environment where users can create and edit pages, making it easy to create documentation, instructions, notes, and other types of content.

Also often used on a project are Microsoft Word, Excell, both documentation formats and attachments for Jira, Confluence. Microsoft Visio is used as a graphic editor to illustrate business processes and diagrams, and for more advanced teams Visual Paradigm, which provides a high level of visualization and project management.

The implementation of an ERP system at any enterprise in Ukraine, as well as throughout the world, is a "revolution" in a commercial organization. Not only all catalogs, components, modules, interfaces, analytics of the existing information system are obtained and revised, but also the structure of the organization, areas of responsibility of departments, and functional responsibilities of employees. The future of the company, not only its efficiency, but also its existence depends on how correctly and efficiently the redesign and implementation of the new information system is carried out.

The main difficulty in implementing an ERP system is the "historically" established processes, the "zoo" of applications and, as a rule, the complete or partial lack of documentation, ranging from user instructions to regulations. A small circle of top-level persons has comprehensive information about the company structure. These employees are often busy and do not express a desire to fully participate in the implementation. And the absence of a business supervisor or lack of control on the part of business process owners leads to design errors and violation of project deadlines.

Often, the implementation of ERP systems is carried out by consulting companies specializing in certain software products. It is better to design and configure an information system with specialists from such a company, who will not only professionally assess the effectiveness of the existing infrastructure, but will also bring internationally recognized practices to the company's work.

Creating an internal company team for the implementation of an ERP system and its fruitful cooperation with the implementing company's team is the most important key to the success of the future project. It is necessary to carefully select the company employees who will take part in the project, determine their degree of involvement and roles. Full allocation of staff to participate in the project is a priority. First of all, it is necessary to find an experienced project manager, functional and technical architect who will participate in the selection of third-party employees with implementation experience, industry knowledge, business analysis and interviewing knowledge.

The first important significance of implementing a new information system is the revision, audit and cleaning of existing directories, which over the years, or even decades, of the company's operation have accumulated a lot of "garbage" and outdated data, which directly affects the system's performance and analytics capabilities.

In second place is putting things in order in architecture. It is no secret that many companies use self-written software, and often the documentation and employees who were at the origins

of its development are no longer available. The unified architecture of the new ERP system, its individual standardized modules with ready-made interface solutions will significantly simplify the administration and security of the company.

General practices that an implemented information system can bring are often sabotaged by staff due to reluctance to master new processes and move away from the "historically" prevailing circumstances in the company. The participation of the company's management in the implementation process is a prerequisite, because a new information system is not a new user interface, it is a new and changed business processes.

The implementation of a new ERP system takes place in parallel with the creation and filling of a project library, which, according to the AIM implementation methodology, contains a description of all the company's business processes, settings, improvements and documentation of changes made. The library of modification specifications will help you understand the coding in the future if you need to make new changes. And a library of user instructions will allow you to avoid gaps in the knowledge transfer to new employees. A description of the settings will help you customize the system more efficiently and avoid desynchronization of modules and loss of efficiency.

The main work when implementing a new IS falls on the business analysts, who must not only describe new business processes in detail and accurately, find "gaps" between the company's existing work processes and the capabilities of the new IS, but also perform reengineering – optimize and rebuild the current ones company practices to improve business efficiency. It is very significance that all company employees from different departments involved in the implementation can correctly describe their functional responsibilities and then understand and approve the new or changed workflow being presented. Often the information is provided in graphical form, so understanding the order of steps, branches, responsible persons, systems, modules and interfaces is necessary for all owners to understand the company's business processes.

Administration and development of changes to the ERP system can be outsourced. But if the company has found and can afford highly qualified technical specialists, then this option will be a priority.

One of the most important project document according to any implementation methodology is the requirement – a functional design document that describes how this or that functionality should look like in the new system. This document is the basis for further development by the technician and writing of user documentation. A functional design document is a common product of the work of the implementation team employee and the company employee who is the owner of the process.

To successfully create a functional design document you must:

- carefully study and describe business requirements;
- develop in detail new or change existing user forms, technical procedures, reports, etc.;
- understand how new changes will affect related processes;
- develop a pool of reports that the employee, the process owner, may need when implementing these changes;
- describe the technical approach that the developer must follow when performing this modification and accept it with the technical architect;

- discuss the design document with analysts and key users;
- obtain approval from the customer.

When preparing documents, all project team members must adhere to the following principles:

- readable, convenient and uniform format;
- compliance with the principles of documentation – maintaining the version of documents, dates of creation/change, indicating the owner, reviewer, approver in all documents;
- it is necessary to avoid nesting other files in Word, Excel, PowerPoint, etc. formats into documents;
- the presentation of information should be understandable both functionally and technically for process owners;
- it is necessary to maintain a register of open/closed questions for each document;
- flowcharts in documents should clearly reflect the client's business flow;
- avoid providing internal intranet URLs, etc. for references;
- all documents must be signed by the client and the head of the company department.

In the process of the research, a survey of specialists in various technical fields was conducted regarding the main problems of IS implementation. The experts were Project Manager, Functional Architect, Technical Architect, Senior Business Analyst, Business Analyst (two), Developer, Commercial Director, Head of Department (two). The results of experts' answers to questionnaire questions are presented in a table 1 and visualized using a bar chart (figure 14). "Expert" is marked as "E" in the table.

Table 1
The results of experts' answers to questions.

Question number	E 1	E 2	E 3	E 4	E 5	E 6	E 7	E 8	E 9	E 10	The sum
Question 1	10	10	10	10	10	10	9	10	10	9	98
Question 2	9	8	9	9	9	9	7	10	7	6	83
Question 3	8	10	9	9	10	10	10	10	8	7	91
Question 4	6	7	6	7	6	6	8	8	9	10	73
Question 5	8	6	8	8	9	9	8	9	8	9	82
Question 6	9	10	10	10	10	10	10	10	7	9	95
Question 7	6	7	8	8	8	7	9	9	10	10	82
Question 8	10	10	10	10	10	10	10	8	9	10	97
Question 9	10	10	10	10	10	10	10	10	10	9	99
Question 10	10	10	10	10	10	10	10	10	8	9	97
Question 11	10	10	10	10	10	10	10	10	9	9	98
Question 12	8	9	10	9	10	9	9	9	6	6	85

According to the method of T. L. Saaty, tables of pairwise comparison of experts' answers to questionnaire questions were constructed and analyzed, Consistency Index (CI) (formula 1), Consistency Ratio (CR) (formula 2) were calculated. According to the tables [6], the RI was determined, for n=10 its value is RI=1.49. The values of CI and CR for each question of the questionnaire are given in the table 2.

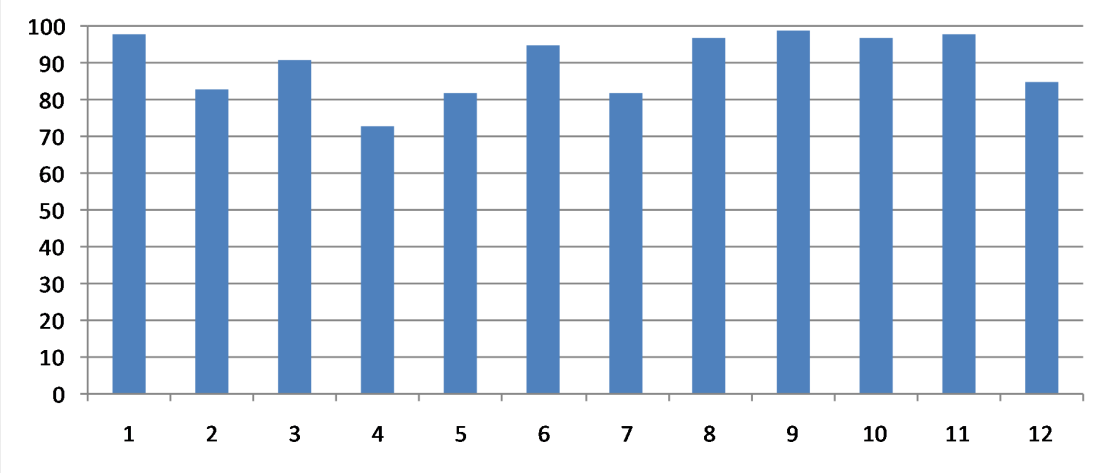


Figure 14: The total number of points awarded by experts to each question.

Table 2
The results of experts' answers to questions.

Indexes	Q 1	Q 2	Q 3	Q 4	Q 5	Q 6	Q 7	Q 8	Q 9	Q 10	Q 11	Q 12
CI	0.00	0.01	0.02	0.01	0.02	0.00	0.02	0.00	0.00	0.00	0.00	0.01
CR	0.00	0.01	0.01	0.01	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.01

According to the values of CI and CR, which are presented in the table 2, it can be concluded that the experts' opinions about each question of the questionnaire coincide, since they do not exceed 0.2 and 0.1, respectively.

4. Conclusion

The development of information systems of commercial enterprises in Ukraine is not much different from global trends, although there is a lag, especially in the field of forecasting and auto orders. Retail companies effort to create a single digital space that unites online and offline sales channels, and e-commerce continues to develop and improve. Analytical systems are beginning to be actively used, but for strategic purposes, although artificial intelligence and machine learning technologies are not yet in demand. Much attention is paid to security and cyber defense, especially during martial law. A cautious trial of cloud platforms is beginning. Increased attention is paid to customer self-service capabilities, focusing on creating more convenient and personalized user interfaces both in online stores and in offline points of sale. There are implementations of robotic systems and process automation in stores and supply chains.

Environmental sustainability is observed – retail companies are increasingly paying attention to environmental issues, in information systems this can manifest itself in accounting and optimization of energy consumption, packaging management, etc.

It can be assumed that the development of Ukrainian retail and the implementation of information systems in this industry will develop dynamically and follow global trends.

This work extends the existing Oracle AIM methodology with practical experience and advice gained from several projects and potential common problems that haunt most ERP system implementation projects in Ukraine. Special attention is paid to the preparation of functional design project documentation.

Experts from different technical fields were surveyed on the main problems of IS implementation. The experiment of analyzing complex solutions using the T. L. Saaty method was performed, and consistency indices were calculated. The analysis of the results of the survey of experts showed that their opinions regarding the answers to the questions of the questionnaire are consistent.

The work can be used as a basis for the preparation of a specialized course on the basics of implementing information systems in commercial enterprises for Professional Junior Bachelors in the field of management and administration. These employees will take part in the description of business processes, study of user documentation and must be oriented in the general structure of the company's information system. They will also need knowledge of the content of the project library and the principles of describing project documentation.

Special attention is paid to the design of project documents, the study and understanding of which is necessary for professional junior specialists, since these employees will be involved in the implementation of the ERP system at the enterprise from the customer's side. Their main duties will be directly related to the description of existing business processes, and the results of implementation depend on the quality of their work.

References

[1] M. S. Alam, M. A. Uddin, Adoption and implementation of enterprise resource planning (ERP): An empirical study, Journal of Management and Research 6 (2019) 1–33. URL: https://www.researchgate.net/publication/334390044.

[2] M. A. Uddin, M. S. Alam, A. A. Mamun, T.-U.-Z. Khan, A. Akter, A Study of the Adoption and Implementation of Enterprise Resource Planning (ERP): Identification of Moderators and Mediator, Journal of Open Innovation: Technology, Market, and Complexity 6 (2020) 2. doi:10.3390/joitmc6010002.

[3] A. G. Chofreh, F. A. Goni, J. J. Klemeš, M. N. Malik, H. H. Khan, Development of guidelines for the implementation of sustainable enterprise resource planning systems, Journal of Cleaner Production 244 (2020) 118655. doi:10.1016/j.jclepro.2019.118655.

[4] S. Menon, Critical challenges in enterprise resource planning (ERP) implementation, International Journal of Business and Management 14 (2019). doi:10.5539/ijbm.v14n7p54.

[5] T. Ivanović, M. Marić, Application of modern Enterprise Resource Planning (ERP) systems in the era of digital transformation, Strategic Management 26 (2021) 28–36. URL: https://www.smjournal.rs/index.php/home/article/view/214.

[6] K. Kulakowski, Understanding the analytic hierarchy process, CRC Press, 2020.

[7] E. Kuzmina, O. Klochko, N. Savina, S. Yaremko, R. Akselrod, C. Strauss, Risk Analysis of

the Company's Activities by Means of Simulation, CEUR Workshop Proceedings 2805 (2020) 162–174. URL: https://ceur-ws.org/Vol-2805/paper12.pdf.

[8] A. Dudek, J. Patalas-Maliszewska, K. Kowalczewska, Automatic Configuration of an Order as an Integral Part of a Cyber-Physical System in a Manufacturing Operating According to Mass-Customisation Strategy, Applied Sciences 13 (2023) 2499. doi:10.3390/app13042499.

[9] P. Rutz, C. Kotthaus, A. F. Pinatti de Carvalho, D. Randall, V. Pipek, The Relevance of KES-Oriented Processes for the Implementation of ERP Systems: Findings From an Empirical Study in German SMEs, Proceedings of the ACM on Human-Computer Interaction 7 (2023) 1–34. doi:10.1145/3610104.

[10] E. Gagnon, Reconceptualizing ERP Integration: A Module Level Perspective, Procedia Computer Science 219 (2023) 596–605. doi:10.1016/j.procs.2023.01.328.

[11] W. H. Salas, Model to improve an ERP implementation based on agile best practice: A Delphi study, Procedia Computer Science 219 (2023) 1785–1792. doi:10.1016/j.procs.2023.01.474.

[12] Oracle Corporation, List of AIM documents, 2011. URL: https://www.oracleerpappsguide.com/2011/03/list-of-aim-documents.html.

Information system module for analysis viral infections data based on machine learning

Nickolay Rudnichenko[1], Vladimir Vychuzhanin[1], Tetiana Otradskya[1] and Igor Petrov[2]

[1]Odesa Polytechnic National University, 1 Shevchenko Ave., Odesa, 65001, Ukraine
[2]National University "Odessa Maritime Academy", 8 Didrichson Str., Odesa, 65029, Ukraine

Abstract

The article presents results of the development information system module for analysis viral infections data. The relevance of the problem of automating the process of analyzing large volumes of data based on the use of intelligent technologies and machine learning methods is considered. The structure of the system has been developed and described, the results of design modeling of the key functionality and capabilities of the system based on the use of the UML language are presented, the basic components and technologies for implementing software are described, allowing for modularity and dynamic expandability of the potential for conducting data analysis research. The process of creating, training and testing the created machine learning models is detailed, the results of assessing the significance of the input features of the collected data set on viral diseases and the obtained values of the error matrices are described. The profiling of the operation process of the created models was carried out, the most productive and efficient of them were determined in terms of the consumption of computing resources and overall accuracy, taking into account their generalization ability.

Keywords

data analysis, data visualization, machine learning, viral infections, information systems development

1. Introduction

In the modern world, information technology (IT) allows for the improvement and acceleration of almost any sphere of human activity, automating key aspects and calculations [1].

Systems and software enable us to eliminate complex tasks and go straight to the results, reducing the time needed to make decisions [2]. Automation of data analysis is one of the main aspects of the practical application of modern IT.

Organizations in various fields focus on collecting and accumulating diverse large volumes of data that need to be gathered, analyzed, and stored in a structured and convenient format for further analysis [3].

The use of modern tools and approaches for data analysis makes it possible not only to research and evaluate statistical indicators but also to perform more complex tasks, including

CS&SE@SW 2023: 6th Workshop for Young Scientists in Computer Science & Software Engineering, February 2, 2024, Kryvyi Rih, Ukraine

nickolay.rud@gmail.com (N. Rudnichenko); vint532@gmail.com (V. Vychuzhanin); tv_61@ukr.net (T. Otradskya); firmn@gmail.com (I. Petrov)
0000-0002-7343-8076 (N. Rudnichenko); 0000-0002-6302-1832 (V. Vychuzhanin); 0000-0002-5808-5647 (T. Otradskya); 0000-0002-8740-6198 (I. Petrov)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

CEUR-WS.org/Vol-3662/paper13.pdf

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

predicting future values using machine learning (ML). In the context of data analysis, one of the current research directions is the issue of viral diseases [4].

Although COVID-19 has become less critical due to regular vaccinations, infection prevention, and the use of preventive measures such as quarantines, viruses have not disappeared from the planet and continue to mutate [5, 6, 7].

Identifying disease patterns, forecasting, and other analyses can help prevent the further spread of the virus. Automation of symptom detection and human body reactions to viruses is essential, as it enables faster modification of effective vaccines and the introduction of various protective measures into the virus transmission environment [8].

The relevance of this work lies in the search for existing methods of data analysis, for their further application to the data visualization and the identification of various patterns based on obtained graphical representations.

The goal of this work is to investigate the dependencies of various indicators in virology using data analysis models, creating a cross-platform information system module for practical purposes.

Assessing and automating data analysis for various diseases can help identify the source of the virus, its symptoms, and prevent its further spread. Collecting such information is a challenging aspect due to the need for access to medical databases or conducting surveys of infected individuals. This complexity reduces the speed of information updates, and the human factor may compromise data reliability, especially when dealing with a significant number of input features [9].

To solve this task effectively, it is important to use the right methods and algorithms during data collection and preprocessing. Specifically, among contemporary approaches in this direction, dimensionality reduction methods are relevant to facilitate the transformation of data into a form suitable for sequential analysis and interpretation of results [10].

The efficiency of data analysis for such data largely depends on the models used and their hyperparameter values [11]. Some of the less deterministic ML models may perform data analysis quickly but produce inaccurate values and lack sufficient generalization capability [12]. On the other hand, more stochastic models may take longer to perform computational operations, but their precision will be higher as a result [13].

For both cases, setting the right attributes for the models and their relationships with each other is a critical factor. This can reduce the likelihood of full-scale quarantines and expedite the development of vaccine modifications [14].

2. Results

2.1. Project structure

During the system development process, UML modeling language was used. The system is built using a client-server architecture and includes a relational database (DB) for storing data from experiments. The developed use case diagram for the system's operation is shown in figure 1.

Within the scope of this diagram, the user can perform the following actions:

- Authenticate in the system using personal data by verifying requests from the database;

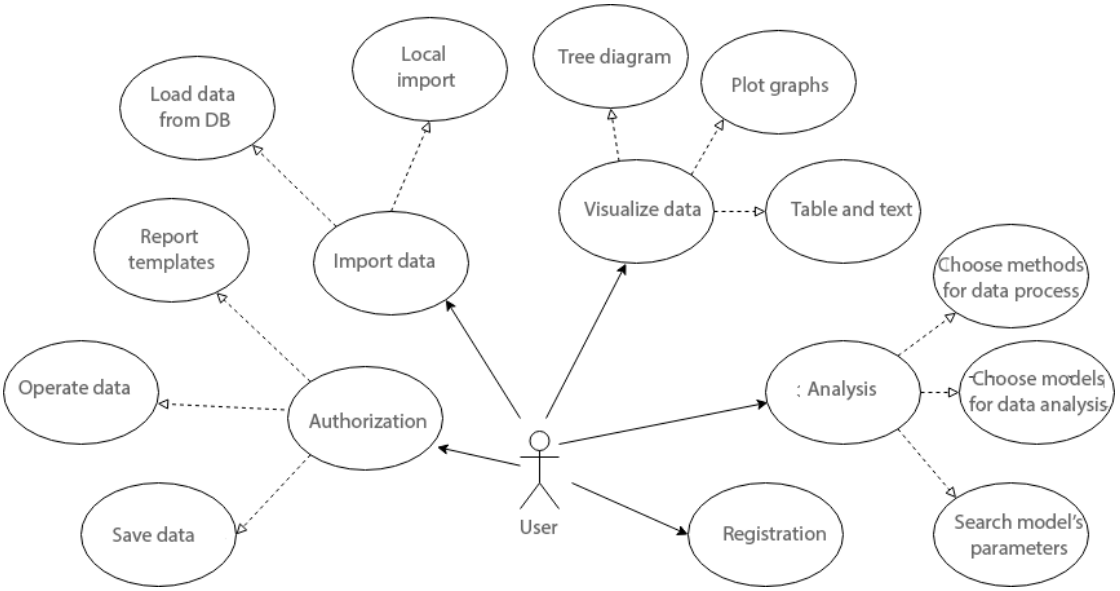


Figure 1: System use-case diagram.

- Import datasets into the system for further analysis;
- Visualize processed data in the form of diagrams, graphs, tables, and textual information;
- Modify arguments and parameters of ML models during their training;
- Register in the system to gain access to data analysis tracking functions.

Let's formalize the main sequence of system operation by applying a sequence diagram for this purpose. For demonstration, we'll take a user who uploads a dataset for conducting exploratory analysis by creating graphical visualizations of static data calculation results (in the form of a small image with statistics and a report with obtained results in a separate file). The user uploads data, selects ML models for analysis, and can set parameter values. The sequence diagram is shown in figure 2.

The user's first action is to access the data analysis model selection form for loading features with parameter editing visualization. After that, they can upload a new dataset in the form of a CSV file or use one of the previously imported data sets serialized to the database. The system user confirms their actions, after which the system dynamically changes the parameter values in the form's interface, depending on which data the user wants to input.

Next, the user fills in all the necessary parameters and data processing algorithms, after which the generated data is sent from the form to the system's handler. Using a REST request, the server-side of the system processes and manages the data to perform further control actions.

In the case of a POST request, the server activates the process of building an ML model based on the specified parameters. The process of using models, in general terms, is as follows:

- Data processing involves receiving input data that will be further processed. Afterward, the system selects software methods for their processing;

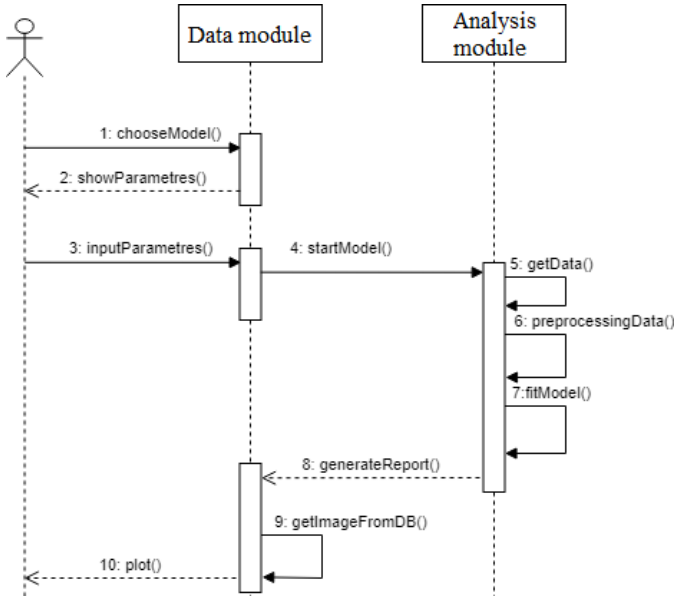


Figure 2: Diagram representing the process of the intelligence analysis of data.

- Creating an instance of an ML model, all imported and processed data are then passed on, and they are processed internally using the fit function;
- Output of a concise report on the evaluation of the performance metrics of the ML model.

After generating the ML model and processing the data, they need to be sent to the system's data visualization module, where, through a GET request, images and other graphical elements specified by the user for the report are displayed. For more detailed planning and formalization of the main relationships between the system components and their structure, it is advisable to construct a diagram of its main components. A generalized component diagram is provided in figure 3.

The main module of the system is flsite. Given that all system modules are developed in Python using the Flask framework, flsite is responsible for the operability and integration of the entire system's functionality. It utilizes the Flask library, allowing the system to be used in the form of a web application, which is convenient due to its compactness and speed.

flsite also contains a set of REST requests that are sent to or received by the system during its operation for rendering authorization pages, user profiles, and model management.

The user profile page contains information about the results of analytical experiments performed by users and datasets uploaded by them. This is necessary for repeating or correcting data analysis.

On the authorization page, the user needs to enter login and password data, as well as a personal token generated by the system after the registration process through a separate modular window. Additionally, for introductory actions in working with the system, the user can choose a limited version in which they don't need to enter authentication data, but they will have restrictions in terms of uploading their data; in other words, they will be able to work

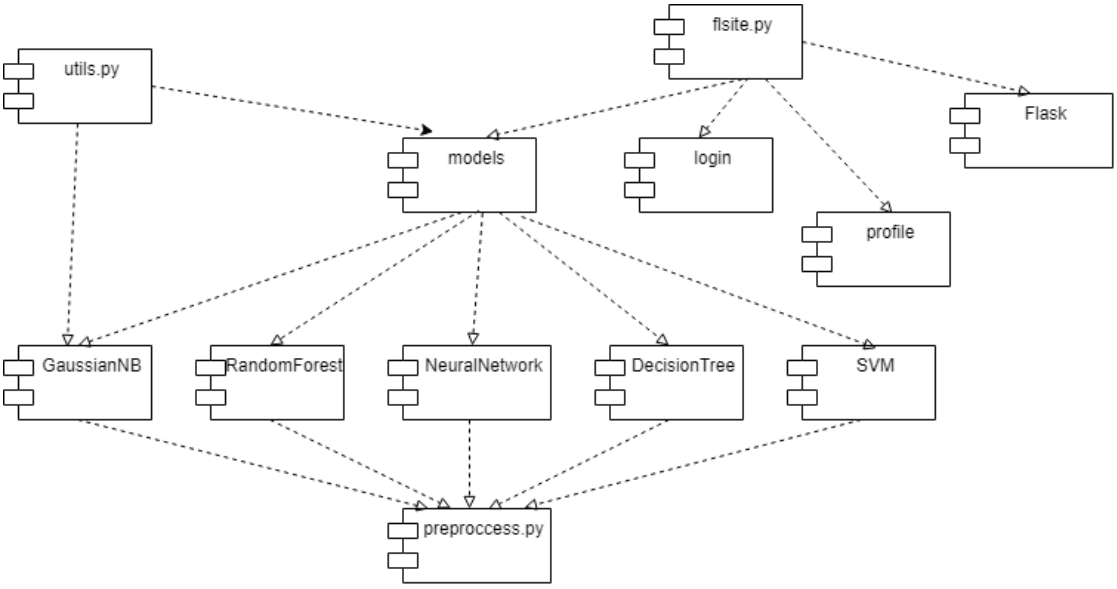


Figure 3: The diagram of the system components in a general form.

with only preinstalled test datasets.

The scikit-learn library is used for the programmatic implementation of ML models due to its convenience, good documentation, and integration with the Python language. The advantage of using this dependency is its support for a significant set of objects for conducting ML model research, performing statistical modeling, including classification, regression, clustering, and dimensionality reduction, through a sequential Python interface. Additionally, libraries such as Pandas, NumPy, SciPy, and Matplotlib were applied for data processing and visualization.

It's worth noting the preprocess.py class, responsible for obtaining input model values and their transformations. This is implemented by integrating functionality from the imbalanced-learn library, which is based on scikit-learn and provides a range of objects for simplified and fast classification work in cases of class imbalance detection.

The utils.py class is responsible for retrieving data from the database, writing them to a file, generating graphical visualizations for reports, transforming data structures and data into matrices of various dimensions, and a variety of other utility functions.

We use SQLite to create the database, which comes bundled with Python3. Its convenience is due to its implementation of support for autonomous and transactional relational mechanisms. The relationship between the database and the system is depicted in figure 4.

First, a database with a set of tables is created separately (without launching the web server). Based on this, request handlers access the construction of database tables, reading and writing information to them. Afterward, the process of creating a general function for establishing a connection to the database and an auxiliary function that will initiate the construction process of a relational database model with the required tables is carried out. At this stage, the open_resource programmatic method is introduced, which opens the 'sq_db.sql' file for reading, located in the project's system working directory. Then, for the open database, the

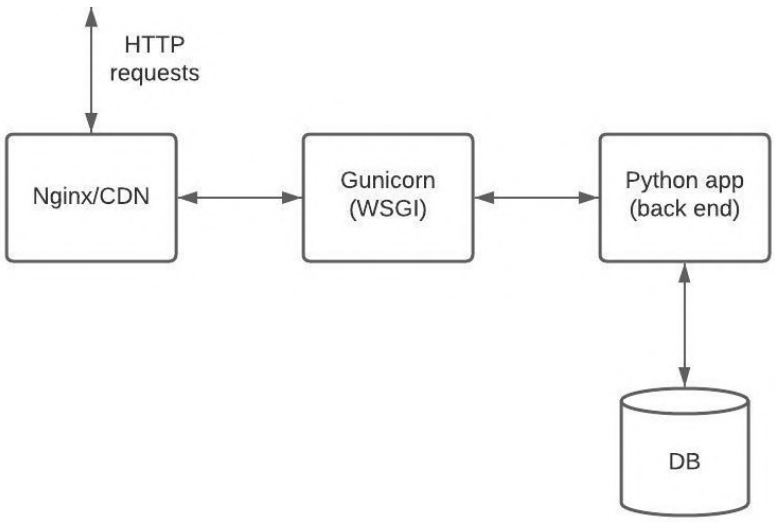


Figure 4: Connection between the system and the database.

script contained in the 'sq_db.sql' file is executed. Finally, the commit method is called to apply the changes to the current database, and the close method closes the established connection.

2.2. ML models implementation

As part of using the system's data analysis module, the following 5 ML models have been programmatically implemented: Gaussian Naive Bayes (GNB), Decision Tree, Random Forest, SVM, and Neural Network (NN) based on a perceptron model. In the system's interface, these models can be assigned various hyperparameter values, allowing clients to input data manually from the web page through corresponding fields. The steps of implementing the data analysis process are as follows:

- Importing input dependencies and libraries.
- Fetching input data into the working environment by downloading them via links.
- Splitting data into training and testing subsets. To perform this step, a cross-validation splitter has been defined, which divides the data into training/test sets according to a specified scheme. Each sample can be assigned to no more than one fold of the test set, as indicated by the user using the test_fold parameter.
- Creating ML model objects.
- Tuning hyperparameter values of ML models using the GridSearchCV method for optimal selection. This is essentially an object implementing a systematic approach to hyperparameter optimization, which involves defining a grid of possible values for each hyperparameter and then exhaustively searching this grid to find the best combination. This approach is called grid search because it involves creating a grid of hyperparameter values and evaluating the algorithm's performance for each combination of values in the grid. It's worth noting that there's no way to know the best hyperparameter values in advance, so ideally, you should try all possible values to find the optimal ones.

- Doing this manually can be time-consuming and resource-intensive, so we use GridSearchCV for automating hyperparameter tuning. The following models have similar code but different parameters within.
- Training models and assessing their accuracy metrics based on the use of methods such as classification_report, taking into account particularities including permutations (permutation_importance). Permutation importance is defined as the reduction in model score when the values of a single feature are randomly shuffled. This procedure breaks the relationship between the feature and the target variable, so a drop in the model's score indicates how much the model depends on that feature. One advantage of this method is that it's model-agnostic and can be calculated many times with different feature shuffles. At this stage, we obtain a finished model for further interaction.

2.3. Data analysis

In order to conduct research, a dataset was created based on information from the European Centre for Disease Prevention and Control regarding the most common COVID-19 symptoms in 2020. The dataset contains information about patients who underwent COVID-19 testing, with a total of 5434 rows and 211 columns. Each row corresponds to a patient's symptoms, where each column represents a symptom. To properly preprocess the data, a dimensionality reduction procedure was performed using PCA. Some features, such as 'had contact with someone with the virus" or "recently traveled", were either removed or combined into more meaningful ones. This allows us to focus on significant symptoms and identify which symptoms indicate early stages of COVID-19.

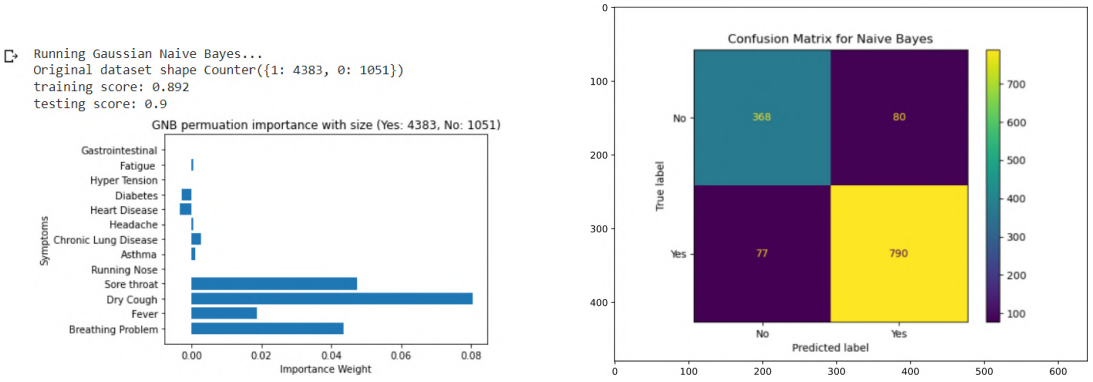
Out of 5434 patients, 81% later tested positive for the virus. To address the dataset's imbalance, the imblearn library was used, which helps minimize the dataset's class imbalance. We chose the resampling method, which aggregates repetitions and reduces data imbalance. The dataset's target variable is expressed in a binary form, with two values: "yes" and "no". Essentially, this allows us to frame the task as binary classification. The data was split into training, validation, and testing sets, with 60%, 20%, and 20% of the data in each part, respectively. Below is an example of the data processing results and model creation with the described datasets.

The results of building the Gaussian Naive Bayes model are shown in figure 5.

As we can see, the top three symptoms for the Naive Bayes model are Dry cough, Sore Throat, and Breathing Problems. Fever was close to the third position, but it's worth noting that its weight is significantly higher than other symptoms. As observed from the matrix, there are 77 False Negatives (FN) and 80 False Positives (FP). Considering that GNB is a basic model, its accuracy is quite high. It's evident that we aim to have a stronger model, and we hope to achieve this with better models ahead. In terms of metrics, the F1 score for negative cases is 0.824, while the F1 score for positive cases is 0.910.

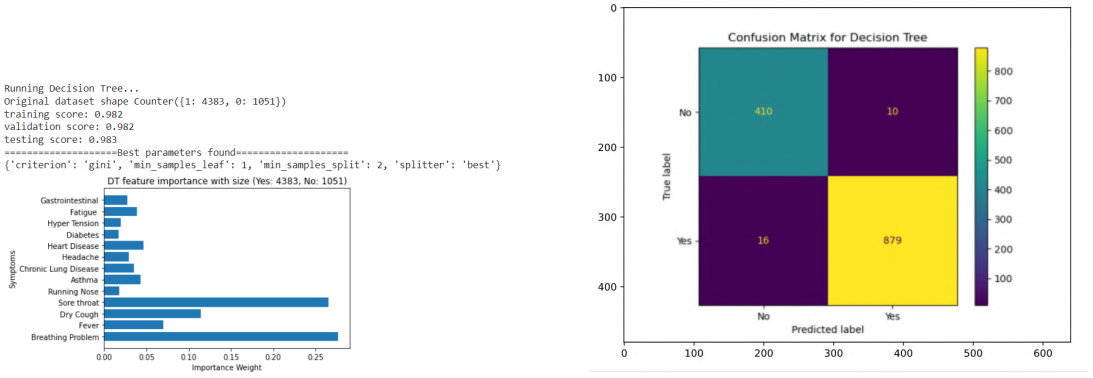
Let's examine the results of the Decision Tree model (figure 6).

As we can see, the top three symptoms for the Decision Tree model are Breathing Problems, Sore Throat, and Dry Cough. Once again, Fever was close to the third position, but, unlike GNB, other symptoms have more weight. In this case, we used the output of the feature_importances from the Decision Tree. There is an interesting shift in the importance of Breathing Problems and Sore Throat. After analyzing the data, we observe an increase in the importance of Sore



(a) (b)

Figure 5: Outputs of GNB model (a) and the corresponding confusion matrix (b).



(a) (b)

Figure 6: Outputs of DT model (a) and the corresponding confusion matrix (b).

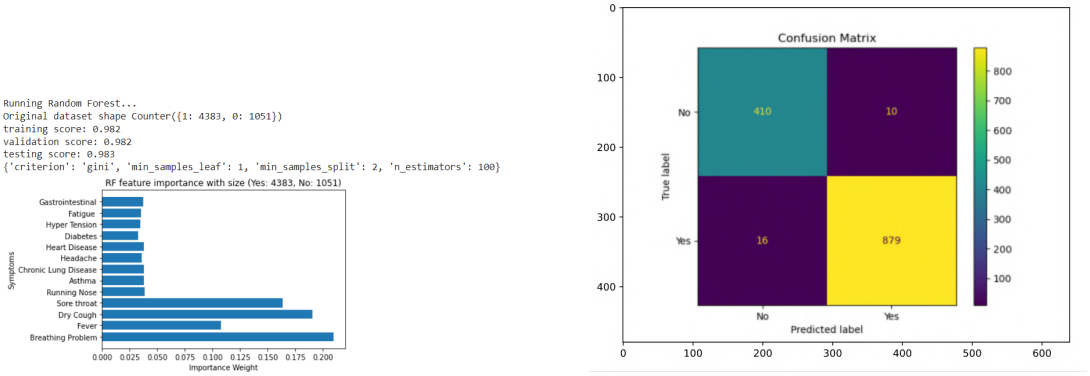
Throat, while Breathing Problems decrease. This suggests that Sore Throat is much more important than Breathing Problems. As we can see, the Decision Tree works much better than GNB. Previously, there were 77 FN and 80 FP, and now there are 16 FN and 10 FP. In terms of metrics, the F1 score for negative cases is 0.969, and the F1 score for positive cases is 0.985.

The most important features for the Decision Tree are Breathing Problems, Sore Throat, and Dry Cough. Currently, the symptoms are consistent with each other in every model we have built. Next, we will conduct research and analyze the Random Forest models (figure 7).

As we can see, the top three symptoms for the scaled Random Forest model are Dry Cough, Breathing Problems, and Sore Throat.

Overall, the Random Forest model produces results similar to the Decision Tree. Previously, there were 16 FN and 10 FP, and we still have 16 FN and 10 FP. In terms of metrics, the F1 score for negative cases is 0.968, while the F1 score for positive cases is 0.985.

The most important features for the Random Forest are Dry Cough, Breathing Problems, and Sore Throat. Currently, the symptoms are consistent in every built and analyzed model.



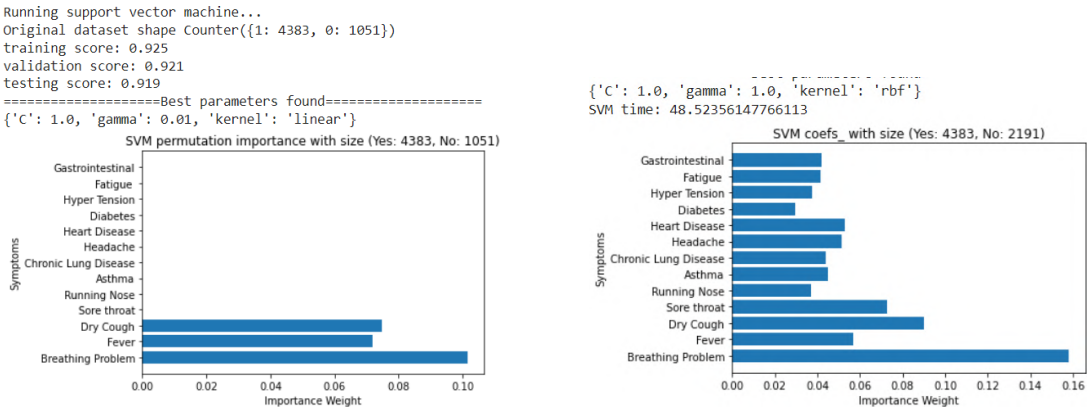
(a) (b)

Figure 7: Outputs of the Random Forest model (a) and the corresponding confusion matrix (b).

Next, let's analyze the performance of the Support Vector Machine (SVM) model. The demonstration of SVM is shown in figure 8. In our case, SVM outputs a hyperplane that separates both classes. The weight coefficients, represented by coef, form a hyperplane orthogonal to the original division boundary. If the hyperplane finds a feature useful for separating the data, the plane will be orthogonal to that axis. Therefore, the coefficient shows how important it was in dividing the two datasets.

We will move on to using more complex kernels (rbf, poly) and use permutation importance to compute feature importance (since coef_ is valid only for linear kernels). As we can see, the top features are Breathing Problems, Dry Cough, and Sore Throat.

In our case, using the RBF kernel works faster and provides a more balanced analysis of feature importance.



(a) (b)

Figure 8: Outputs of the SVM model using linear kernal (a), and using RBF kernel (b).

Let's conduct an investigation of the created neural network model (figure 9). For the neural

network, we compare functions using permutation importance.

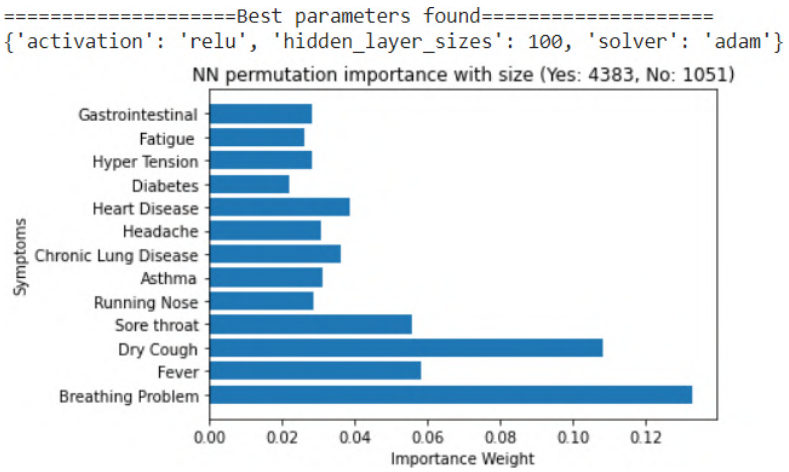


Figure 9: Outputs of the Neural Network.

From the parameter grid, we used a validation set to find the best combination. For the activation function, we used ReLU, the number of hidden layers is 80, and we utilized the Adam optimizer for better performance.

We wanted to understand whether our network prefers having one larger bank of hidden layers or several smaller banks of hidden layers. The artificial neural network model achieved the most significant results with only 6 cases of TN errors. The most important features are "breathing problems", "dry cough", and "sore throat", depending on the dataset used.

Based on the results obtained, let's analyze the performance speed of all the ML models discussed above. The speed of the two neural networks (ReLU and Adam) is presented through profiling in figure 10. In this case, we can see that GNB and DT have the minimum operating speed. RF, due to its complexity, operates more slowly. SVM and NN1 work significantly faster, even with input value analysis.

Therefore, the developed module is sufficiently functional and allows for the exploration of datasets to uncover hidden patterns within the data.

3. Summary

As a result of this research, a software module for the system has been developed, capable of performing data preprocessing and analysis steps related to human virus-related illnesses using machine learning methods. The following evaluation criteria for the models were determined: speed, accuracy, predictions, best hyperparameters for the models, error matrix assessments.

The research findings have established that the most effective models are artificial neural networks, but decision trees also showed one of the best results, considering that this model is based on a basic algorithm. The study also revealed the varying speeds of different models, with a significant workload dependence on the chosen hyperparameter values. Additionally, data inaccuracies can complicate the process of determining the best model.

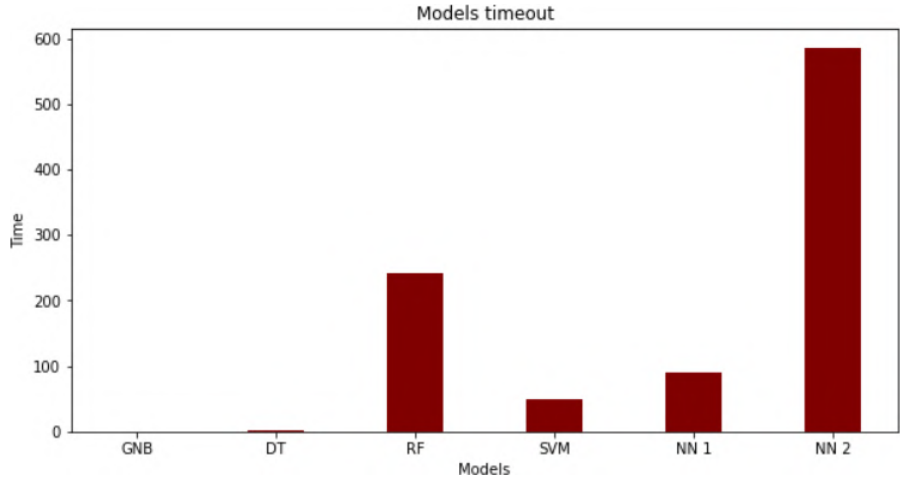


Figure 10: Comparison of the time performance (cost) of the created models.

Therefore, it can be concluded that the best model for working with the data collected during the development of the system is the neural network. However, its speed is significantly lower than that of the decision tree model. Therefore, future research in this area could involve exploring more efficient models and forming a more extensive dataset for a more balanced analysis.

References

[1] C. Peng, Digital Inclusive Finance Data Mining and Model-Driven Analysis of the Impact of Urban-Rural Income Gap, Wireless Communications and Mobile Computing 7 (2022) 1–8. doi:10.1155/2022/5820145.

[2] N. Rudnichenko, V. Vychuzhanin, I. Petrov, D. Shibaev, Decision Support System for the Machine Learning Methods Selection in Big Data Mining, CEUR Workshop Proceedings 2608 (2020) 872–885. URL: https://ceur-ws.org/Vol-2608/paper65.pdf.

[3] I. M. Shpinareva, A. A. Yakushina, L. A. Voloshchuk, N. D. Rudnichenko, Detection and classification of network attacks using the deep neural network cascade, Herald of Advanced Information Technology 4 (2021) 244–254. doi:10.15276/hait.03.2021.4.

[4] J. Andry, H. Tannady, G. Rembulan, D. Dinata, Analysis of the Omicron virus cases using data mining methods in rapid miner applications, Microbes and Infectious Diseases 4 (2023). doi:10.21608/mid.2023.194619.1469.

[5] H. Wijaya, S. Syaifudin, T. Siswanto, Visualization of corona virus disease 2019 deoxyribonucleic acid data analysis, AIP Conference Proceedings 2659 (2022) 090005. doi:10.1063/5.0118893.

[6] A. D. Kubegenova, E. S. Kubegenov, Z. M. Gumarova, G. A. Kamalova, G. M. Zhazykbaeva, Using Data Mining Technology in Monitoring and Modeling the Epidemiological Situation of the Human Immunodeficiency Virus in Kazakhstan, in: A. Gibadullin (Ed.), Information

Technologies and Intelligent Decision Making Systems, Springer Nature Switzerland, Cham, 2022, pp. 57–65. doi:10.1007/978-3-031-21340-3_6.

[7] P. Srinivas, D. Bhattacharyya, D. MidhunChakkaravarthy, Prediction of Swine Flu (H1N1) Virus Using Data Mining and Convolutional Neural Network Techniques, in: A. Kumar, G. Ghinea, S. Merugu, T. Hashimoto (Eds.), Proceedings of the International Conference on Cognitive and Intelligent Computing, Springer Nature Singapore, Singapore, 2023, pp. 557–573. doi:10.1007/978-981-19-2358-6_51.

[8] H. Gunawan, V. Purwayoga, Data mining menggunakan algoritma k-means clustering untuk mengetahui potensi penyebaran virus corona di kota cirebon, Jurnal Sisfokom (Sistem Informasi dan Komputer) 11 (2022) 1–8. doi:10.32736/sisfokom.v11i1.1316.

[9] V. Dev, P. Singh, Effect of Corona Virus on Multi-Disease Patients using Association Rule Mining, 2023. URL: https://www.researchgate.net/publication/372312135.

[10] E. Luna-Ramírez, J. Soria-Cruz, I. Castillo-Zúñiga, J. I. López-Veyna, COVID-19 Social Lethality Characterization in some Regions of Mexico through the Pandemic Years Using Data Mining, in: Y. P. Rybarczyk (Ed.), Research Advances in Data Mining Techniques and Applications, IntechOpen, Rijeka, 2023. doi:10.5772/intechopen.113261.

[11] A. Y. O. Allmuttar, S. K. D. Alkhafaji, Using data mining techniques deep analysis and theoretical investigation of COVID-19 pandemic, Measurement: Sensors 27 (2023) 100747. doi:10.1016/j.measen.2023.100747.

[12] M. Pandey, Machine Learning, International Journal for Research in Applied Science and Engineering Technology 11 (2023) 864–869. doi:10.22214/ijraset.2023.55224.

[13] D. P. F. Möller, Machine Learning and Deep Learning, in: Guide to Cybersecurity in Digital Transformation: Trends, Methods, Technologies, Applications and Best Practices, Springer Nature Switzerland, Cham, 2023, pp. 347–384. doi:10.1007/978-3-031-26845-8_8.

[14] R. Thakur, P. Panse, P. Bhanarkar, Machine Learning and Deep Learning Techniques, in: U. N. Dulhare, E. H. Houssein (Eds.), Machine Learning and Metaheuristics: Methods and Analysis, Springer Nature Singapore, Singapore, 2023, pp. 235–253. doi:10.1007/978-981-99-6645-5_11.

Designing a cross-platform user-friendly transport company application

Maksym Y. Salohub, Olena H. Rybalchenko and Svitlana V. Bilashenko

Kryvyi Rih National University, 11 Vitalii Matusevych Str., Kryvyi Rih, 50027, Ukraine

Abstract

This paper delves into the pivotal role of cross-platform development and the significance of user experience (UX) and user interface (UI) research in the creation of an application for a transportation company. User experience involves studying user behaviors, preferences and interactions, while user interface focuses on the design elements that facilitate these interactions. The research methodology includes a user experience research, including a user survey and analysis of analogous applications. The findings from this study are instrumental in researching the visual aspect of the application, aiding in the identification of design solutions and the creation of interface layouts. The paper analyzes technologies that facilitate cross-platform development. By drawing conclusions from the user experience research, interface design development and an examination of cross-platform development tools, the study culminates in the development of a convenient and scalable application tailored for both users and drivers of the transportation service.

Keywords

cross-platform application, software, React Native, mobile, transport, interface, UI/UX, nativity

1. Introduction

In the era of information technology and digitalization of all spheres of life, an important aspect of business is the availability of proprietary software. Achieving cross-platform applications can be done in two ways: using native approaches for each platform or employing technologies for a unified code base. A single code base reduces development costs and enables smaller companies to afford cross-platform applications, increasing demand.

For transportation companies, the ability to offer apps is usually limited to industry giants, leaving smaller regional providers with fewer options. Therefore, exploring the creation of affordable cross-platform applications is crucial for the modern market.

Given the active interaction between transportation services, users, and drivers, conducting user experience research and user interface analysis is essential. This ensures the development of a modern, high-quality design that is minimalistic and non-distracting for drivers. Addressing

CS&SE@SW 2023: 6th Workshop for Young Scientists in Computer Science & Software Engineering, February 2, 2024, Kryvyi Rih, Ukraine

taranmak21@gmail.com (M. Y. Salohub); rybalchenko@knu.edu.ua (O. H. Rybalchenko); bilashenko.s@knu.edu.ua (S. V. Bilashenko)

http://mpz.knu.edu.ua/vikladachi/olena-rybalchenko (O. H. Rybalchenko); http://mpz.knu.edu.ua/vikladachi/Svetlana_Bilashenko (S. V. Bilashenko)

0009-0007-9696-3766 (M. Y. Salohub); 0000-0001-8691-5401 (O. H. Rybalchenko); 0000-0002-4331-7425 (S. V. Bilashenko)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

CEUR-WS.org/Vol-3662/paper18.pdf

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

UI/UX aspects before full development is crucial to optimize resources and enhance the overall application experience.

Another problem with cross-platform development can be that it is difficult to adhere to a single style, because some tools can cause a conflict while displaying interface elements on different platforms. Therefore, after creating a full-fledged design system, it is necessary to select elements that may cause a display conflict and make them custom so that the systems generate them instead of taking them from the ready-made tools. Thus, the developed application will be ready to be deployed to the AppStore and PlayMarket without any problems with platform compatibility [1].

2. Researching approaches to cross-platform development

The principle of a cross-platform application is generally to use a single code base that can be executed on any platform without significant modifications or without any modifications at all. However, cross-platform is a complex thing to implement and has several options for creating it [2].

Some applications use compilation to native code, which allows the application to run as native on a particular platform. Native elements are those that are unique to a particular development platform, even though they have the same name. So, a button on two platforms is still a button, but it may look different. But applications that create the so-called "bridge" can interact with both their own program code and the program code of the platform environment on which they were launched [3].

Some applications use web technologies to display the application. Thus, the use of frameworks in software systems whose work is based on creating mappings in the web environment allows the application to avoid cases of different interface appearance when running on different platforms. The advantage of such systems is that only one developer who has created a responsive website is enough. But for the average user, such applications have a huge drawback: they must always load a web page and therefore are limited to working only on devices that have stable Internet access and a built-in browser.

It is also possible to distinguish the way the program works when it does not have access to the native elements of the platform. Such applications use their own graphics tools, thus gaining access to the device's GPU. They use it to display the same interface elements regardless of the platform. But, like every other type of cross-platform application, they have their drawbacks, which in certain situations can be quite significant for the user. For example, they put more strain on weaker devices.

The larger the city and the more places where tourists visit – the more the application must support a large number of requests to the network. Therefore, it is extremely important to design high-quality work with network requests for both iOS and Android systems [4].

3. UX research

The UX research was divided into two parts: competitor analysis and target audience analysis using a questionnaire [5]. To analyze analogs, was selected the applications that also use cross-

platform technologies. These are applications from Bolt, Grab and Didi Rider. The analysis results were obtained according to visual analysis and analysis of users reviews (table 1).

Table 1
Results of analogs analyze.

Parameter	Bolt	Grab	DiDi Rider
Popularity	Widely popular	Some popular	Not so popular as analogs
Multiple services in one application	Scooters and passenger transportation	Passenger transportation and food delivery	Only passenger transportation
Cash and non-cash payments support	Widely support	Widely support	No support
Geographical availability	Limited	Limited	Limited
User Experience quality	High quality of UX	Some UX issues	A lot of UX issues
Service issues	Rare issues	Rare issues	Often issues
Ride request edit availability	Can't edit	Can't edit	Can't edit

The next step of UX research is to survey users. The most convenient way to do this is to use Google Forms with a set of questions [6]. At this stage, 30 people were interviewed and the results of the survey were compiled. For a simpler understanding of the results, separate charts and tables were created for the most important points of the survey. Thus, the most voluminous, but important chart was the chart of the type of people who use the app.

Next, the most important part of the survey is which app features are important to users. Thus, for further work, it becomes possible to easily identify the main functionality that is really relevant for the modern market of transport companies. The results from the chart were transformed into a table for easier data processing (table 2).

Table 2
Important functions for users.

Question	Count	Percentage
Driver choosing availiability	16	53.3%
Showing driver photo while choosing driver	19	63.3%
Automatic drivers filtration by geolocation	23	76.7%
Chat with driver only when active trip	10	33.3%
Ending of trip when user and driver agreement both	13	43.3%
Panic button	23	76.7%
Trips archive	17	56.7%
Trip review	20	66.7%
The ability for drivers to take orders at any free time	1	3.3%
The function of confirming the arrival of a taxi only with my permission	1	3.3%

According to this chart, the two most important features for users are the presence of a panic button during the trip to ensure their safety and automatic filtering of drivers by trip details (for example, by selecting the "with child" tag for a trip, drivers without a child seat will not even see this order). As slightly less important features, respondents noted the ability to rate a ride, see a driver's photo when choosing a driver, the ability to choose a driver and save a ride history. Also, almost 43.3% of users consider it necessary to be able to end the ride only upon their confirmation, so the driver cannot disable the transportation on his/her own. The least important features are the activity of chatting with the driver only during the trip itself, the ability for drivers to take orders at any free time and one answer was added by the user personally, which sounds like "the function of confirming the arrival of a taxi only with my permission". Thus, the most basic needs of application users in terms of functionality that need to be implemented first was identified.

Equally important is not only the implementation of the functionality, but also what exactly will be displayed on the screen with a higher priority. To determine this, the question "what is paid attention to when choosing a transport" was separately highlighted and a diagram was created based on the answers. The results of diagram were transformed into a table (table 3).

Table 3
Focus of the attention of users when choosing a transport.

Question	Count	Percentage
Cost	24	85.7%
Car brand	5	17.9%
Driving experience	7	25%
Company reputation	15	53.6%
Quality of service	20	71.4%
Payment convenience	19	67.9%
Driver rating	1	3.6%

Let's summarize the results of the UX research. The app's users are active citizens who value their time and comfort. They use smartphones with internet access, which makes them ideal users. They also have different socio-economic statuses, including students, office workers, entrepreneurs, tourists and drivers looking for additional income.

4. UI analysis, creating interface prototypes and design system

The development of the UI part of the application can also be divided into several parts – creating low-detail mockups (wireframing), graphic design and designing full-fledged design mockups. Wireframing helps to visualize the concept of the application and provides an opportunity to quickly check and improve the structure of the product before moving on to more detailed design and development. After low-detail prototypes are created, they are refined into highly detailed mockups, where colors, textures, graphics and other elements are added to increase the level of detail and prepare the mockup for use by the developer when building the app [7, 8].

First, was created a prototype of the login screen, because this is the first thing the user sees,

but it is more important to consider the next screen after the login. This is the home page screen. It is divided into two versions, because the user can be either a passenger or a driver. Since they are still quite similar, let's take a look at the passenger screen (figure 1).

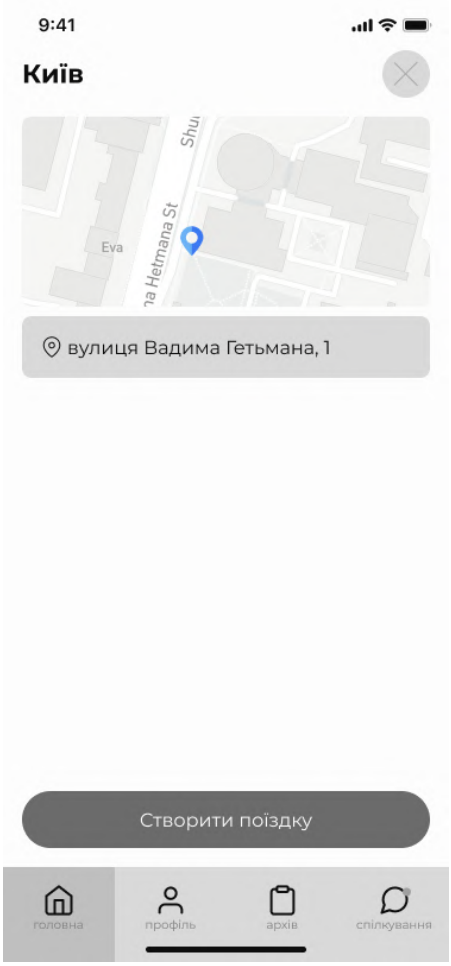


Figure 1: Prototype of the main screen of passenger.

It was also useful to create a prototype of the profile screen to understand the information architecture of the application (figure 2).

These were all examples of creating low-detail interface layouts. Their task is to structure the data on navigation in the application and the general arrangement of elements.

The next stage of UI modeling is graphic design. When the information architecture of the application is already known, it becomes important to select the right icons that will fully express the semantic meaning that was embedded in them. That's why status icons were identified as part of the design system (figure 3).

As you can see, some of these icons are also colored – this is necessary to further emphasize to the user what exactly has happened in the application at the moment. Thus, additional

Figure 2: Prototype of main profile screen.

Figure 3: Icons for informing about processes in the application.

semantic coloring is provided for the images.

So, when the low-detail mockups and the graphic part of the design are ready, you can move on to developing high-detail mockups. Since, according to the results of the UX analysis, it is clear that the users of the application are a wide range of people, a blue was chosen as the main color. It is associated with calmness and reliability, which is essential for a transportation company (figure 4, 5).

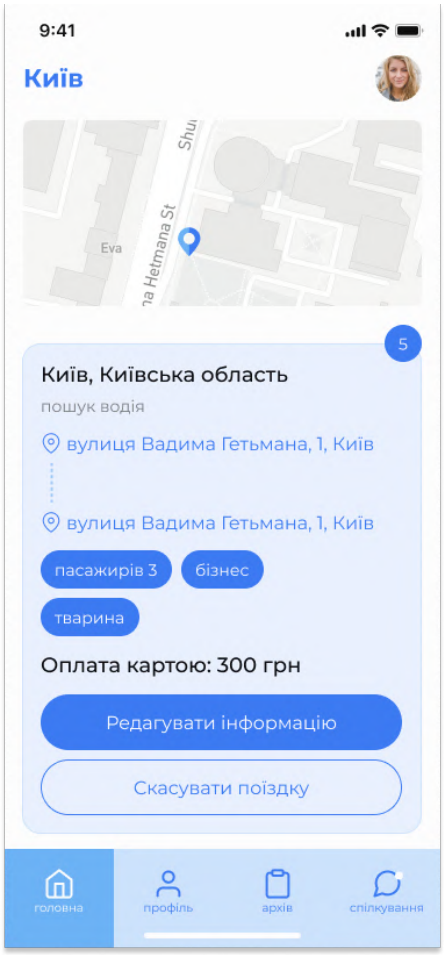


Figure 4: Main screen of the user.

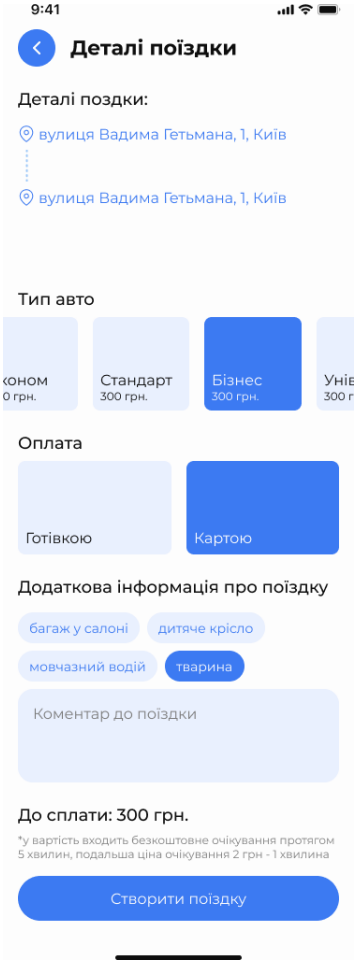


Figure 5: Trip confirmation screen.

A grouping of screens based on a specific set of functionality is called a flow. For example, this is how a chat flow looks like (figure 6).

As a result of the work, we have a completely ready-made design system that can be used for software development, knowing that it will be useful and convenient for each user.

5. System development

To implement the application, JavaScript was chosen to avoid switching between languages when creating different parts of the project. Thus, the Express.js framework is used for the backend part and React Native is used for the application itself.

React Native is used to develop mobile applications. The main advantage of React Native is that this technology allows you to write code once and then use it to create applications for both iOS and Android. This significantly reduces the development time and eliminates the need to write two separate programs for different platforms [9].

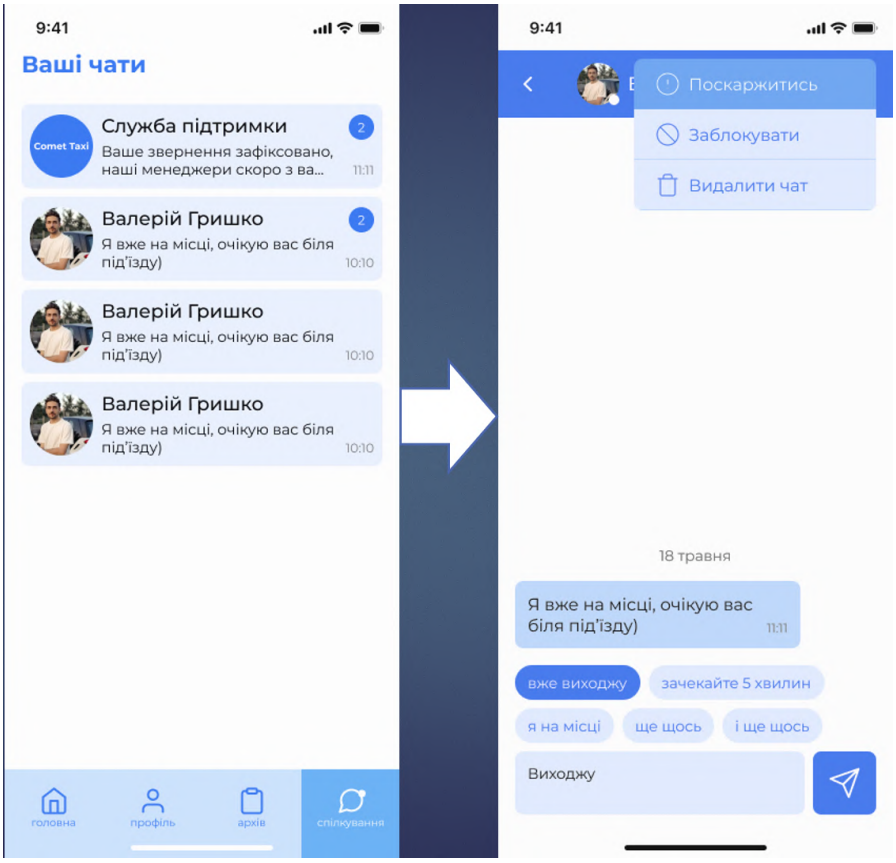


Figure 6: Chats flow.

Express.js is used to develop the server side of the system. It is a flexible and powerful framework that allows you to easily create web applications and APIs. Express.js integrates well with MongoDB and other technologies used in the project [10].

Another feature of the project is the use of Expo CLI. One of the most important features of Expo is the ability to develop mobile applications on the React Native platform. With Expo, you can easily build, test and publish your React Native apps without having to set up complex development tool configurations. Testing is done by scanning a QR code from the IDE terminal (figure 7) [11].

This project uses MongoDB to store data. It is one of the most popular NoSQL database management systems. MongoDB is a document-oriented database that uses a JSON-like structure to store data rather than working with tables in the usual sense of databases. Features of MongoDB include data structure flexibility, high performance, horizontal scaling and the ability to work with large amounts of data. The database is created using MongoDB Atlas, a MongoDB cloud service [12].

A special collection of icons called Feather Icon was used in the application. With the help of this service, a collection of icons was created that are used to improve the semantic interaction

Figure 7: Terminal interface for testing.

of users with the application. React Native technology does not support easy interaction with SVG icons. But this is necessary, because only this image format scales to any screen size without changing the quality. Therefore, to display such icons in the application, additional libraries react-native-svg and react-native-feather are used. This way, a contact is established between the application and the cloud service, which is the repository of the necessary icons in SVG format.

There is a separate algorithm for saving images on the server. Due to the use of MongoDb, it was decided to encode the images transferred to the server in base64 using the multer auxiliary library.

The basic encryption code looks like this:

```
const multer = require('multer');
const storage = multer.diskStorage({
    destination: function (req, file, cb) {
        cb(null, './uploads/');
    },
    filename: function (req, file, cb) {
        cb(null, new Date().toISOString() + file.originalname);
    },
});
const upload = multer({ storage: storage });
module.exports = upload;
```

The code of the auxiliary file used for intermediate image encoding looks like this:

```
const multer = require('multer');
const storage = multer.memoryStorage();
const upload = multer({
    storage: storage,
    limits: {
        fileSize: 1024 * 1024 * 2,
    },
});
module.exports = upload.single('image');
```

It is important to note that this file size was not chosen by chance. It was important to save as much space as possible in the database without loading the server with long image processing. Therefore, the image will be successfully saved only if its size is 1024*1024 or less. At the same time, the maximum image size is limited to 2 MB.

6. Conclusions

This project was a fascinating challenge that allowed us to delve deeper into the world of developing and researching the benefits of cross-platform applications for transportation companies. One of the key points of focus was to create an application that not only meets the needs of users, but also has a convenient and efficient interface for both parties – both passengers and drivers.

A number of technologies were used to achieve this goal. React Native allowed us to efficiently implement the visual particle, ensuring its cross-platform compatibility. Express.js was used to create a powerful backend and MongoDb served as a reliable database.

Analysing of the subject area, was identified the main problems of interaction between passengers and drivers. This insight allowed us to clearly define the project's goals and objectives, focusing our efforts on solving specific problems.

Target audience research confirmed that drivers and passengers using modern mobile technologies are the main users of the app. The results of the survey emphasized the high demand for such services, which became the basis for further improvement of the app.

Overall, this project not only demonstrated the effectiveness of using modern development methodologies and tools, but also emphasized the flexibility and adaptability of the application to changes in the passenger transportation market. The result is a high-quality mobile application that is ready to quickly adapt to any conditions and user requirements.

References

[1] N. V. Ichanska, S. I. Ulko, Osnovni aspekty stvorennia mobilnykh dodatkiv ta vybir instrumentiv yikh rozrobky [The main aspects of creating mobile applications and the choice of tools for their development], Systemi upravlinnia, navigatsii ta zv'iazku 59 (2020) 74–78. doi:10.26906/SUNZ.2020.1.074.

[2] M. Martinez, S. Lecomte, Towards the quality improvement of cross-platform mobile applications, 2017. URL: http://arxiv.org/abs/1701.06767.
[3] B. Eisenman, Learning React Native: Building Native Mobile Apps with JavaScript, 2nd ed., O'Reilly Media, Inc., 2017.
[4] A. Noulas, V. Salnikov, D. Hristova, C. Mascolo, R. Lambiotte, Developing and Deploying a Taxi Price Comparison Mobile App in the Wild: Insights and Challenges, 2017. URL: http://arxiv.org/abs/1701.04208.
[5] R. L. Veal, How to Conduct User Experience Research Like a Professional, 2021. URL: https://careerfoundry.com/en/blog/ux-design/how-to-conduct-user-experience-research-like-a-professional/.
[6] M. Y. Salohub, Opytuvannia dlia vivchennia tsilovoi auditorii dlia stvorennia dodatku sluzhby taksi [Survey to study the target audience for creating a taxi service app], 2022. URL: https://docs.google.com/forms/d/e/1FAIpQLSerQrKhXxOJ7s1rqLY5fAyWI_9HHUgLS4HVPGarZb6k9Gc0Ag/viewform.
[7] J. Yablonski, Laws of UX, 1st ed., O'Reilly Media, Inc., 2020.
[8] I. D. Foundation, What is Wireframing: The complete guide, 2023. URL: https://www.interaction-design.org/literature/topics/wireframing.
[9] Meta, React Native: Get Started, 2023. URL: https://reactnative.dev/.
[10] Expressjs.com, Express js backend documentation, 2023. URL: https://expressjs.com/.
[11] Expo.com, Expo developer documentation, 2023. URL: https://expo.dev/.
[12] Mongodb.com, MongoDB documentation, 2023. URL: https://www.mongodb.com/.

Research of the route planning algorithms on the example of a drone delivery system software development

Yevhen L. Turchyk[1], Milana V. Puzino[2], Olena H. Rybalchenko[1] and Svitlana V. Bilashenko[1]

[1]Kryvyi Rih National University, 11 Vitalii Matusevych Str., Kryvyi Rih, 50027, Ukraine
[2]Lviv Polytechnic National University, 12 Stepana Bandery Str., Lviv, 79000, Ukraine

Abstract

The paper analyzes the existing drone delivery systems all around the world. Different route building algorithms are analyzed for navigating the drones through the cities, advantages and disadvantages of all the approaches are highlighted. Requirements for the system are defined that must provide quick and convenient operation; the system was planned and developed. It was concluded that the designed system has a great potential for real usage and further development.

Keywords

drone delivery, UAV, path finding, route building, machine learning

1. Introduction

In the modern world, there is a noticeable acceleration of the pace of life in large cities. The efficiency of businesses and the quality of life for individuals depend directly on well-established logistics. This issue is particularly pronounced in "last-mile" delivery, where the transportation of goods is influenced by various external factors that impact its speed and effectiveness. First and foremost, human labor involved in product delivery is limited by physical and psychological aspects, and the increasing demand for speed may exceed the capabilities of personnel. The involvement of significant resources in the delivery process can result in increased delivery costs, subsequently raising the prices of goods and services for both businesses and end consumers.

Drone delivery can address these challenges. The use of drones in delivery can reduce the dependency on human labor. Unmanned aerial vehicles (UAVs) can operate around the clock without rest, providing fast and precise product delivery when using advanced navigation and route building algorithms [1].

CS&SE@SW 2023: 6th Workshop for Young Scientists in Computer Science & Software Engineering, February 2, 2024, Kryvyi Rih, Ukraine

itsjonny5757@gmail.com (Y. L. Turchyk); milana.puzino.mpzip.2022@lpnu.ua (M. V. Puzino); rybalchenko@knu.edu.ua (O. H. Rybalchenko); bilashenko.s@knu.edu.ua (S. V. Bilashenko)
http://mpz.knu.edu.ua/vikladachi/olena-rybalchenko (O. H. Rybalchenko); http://mpz.knu.edu.ua/vikladachi/Svetlana_Bilashenko (S. V. Bilashenko)
0009-0006-3254-9411 (Y. L. Turchyk); 0009-0007-9450-630X (M. V. Puzino); 0000-0001-8691-5401 (O. H. Rybalchenko); 0000-0002-4331-7425 (S. V. Bilashenko)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

CEUR-WS.org/Vol-3662/paper21.pdf
CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

Therefore, the primary idea of this work is to study the path finding algorithms and develop a software for automated aerial drone delivery to address the issue of "same-day" delivery between different city branches and accelerate it. To achieve this, a comprehensive software solution is proposed to implement a similar service and explore the capabilities of route planning algorithms for the automatic operation of UAVs.

2. Review of the subject area and existing solutions analysis for the development of drone delivery system

2.1. Existing systems analysis

Currently, there are relatively few existing and fully operational drone delivery analogs on both the Ukrainian and global markets. Most of the available services are either in the testing or development stages or operate within limited geographical areas. Additionally, the majority of these services are oriented towards "last-mile" delivery, which restricts users from utilizing the service for non-commercial purposes. Let's examine some of the most well-known analogs within the mentioned category.

2.1.1. Amazon Prime Air

Amazon Prime Air is an aerial drone delivery project developed by Amazon since 2013, aiming to provide rapid delivery of packages to recipients within 30 minutes [2].

Currently, the system is in the testing phase, conducted in the city of Lockford, California, USA. It is expected that based on feedback from local residents using the delivery service, the most problematic aspects will be identified and addressed. Through testing with a wide variety of cargo sizes and weights, the reliability and durability of the project's technical equipment have been verified.

The following advantages of this service can be highlighted:

- According to reports from Amazon Prime Air project specialists, unique software for drones has been developed, allowing UAVs to safely detect and avoid potential obstacles, making the point-to-point flight process more reliable.
- Amazon Prime Air drones have a high payload capacity from the outset.

However, it is worth noting the drawbacks of such a system:

- The process of delivering cargo to the recipient involves dropping it from a specified height onto the backyard of a private house, limiting the potential user base to those with suitable delivery locations. This approach may not guarantee the safe delivery of potentially fragile cargo.
- This system is planned to be used exclusively for delivering goods from the Amazon store, which narrows down the pool of potential users.

It is expected that users will have the option to order drone delivery of selected items through the Amazon online store, and as such, this service will not have its own separate user software but will be integrated into the existing services of the company.

2.1.2. Starship Technologies

Starship Technologies is an Estonian startup (later becoming a company) initiated in 2014, addressing the "last-mile" delivery problem using ground-based drones [3].

This service operates in London, Tallinn, Düsseldorf, Hamburg, Bern, and in some cities in the United States, such as Washington, D.C., and Mountain View, California. Among the advantages of this startup, the following points are noteworthy:

- The drones are fully autonomous, allowing them to independently locate and load the required product into their cargo compartment.
- Delivery is secure for the recipient since receiving an order is only possible after entering a personal security code.
- In case of navigation issues with the drone, the system provides remote control by a human pilot.

However, there are some limitations to the startup:

- Deliveries are only made within a 5-kilometer radius.
- The maximum drone speed is determined by the quality of the terrain and does not exceed 6.5 km/h, which is nearly equivalent to human walking speed.
- The project is exclusively oriented towards delivering food weighing up to 9 kg.

This described service has a dedicated user application through which the ordering, payment, package tracking, and other processes are conducted.

2.1.3. Zipline

Zipline is an American project involved in the manufacturing and delivery of air drones/aircraft [4]. The main concept of the company is to address the issue of delivering cargo to hard-to-reach locations.

Currently, the Zipline service is available in Rwanda, Ghana, Nigeria, Japan, and the United States. Additionally, it is expected that the company's services will soon become available in Côte d'Ivoire and Kenya. Furthermore, the Ministry of Health of Ukraine has announced negotiations regarding a potential partnership with the company.

Notable advantages of the Zipline service include:

- Ensuring high delivery speed, even for long distances, thanks to the mobility of the drone-aircraft (UAV speed reaches 101 km/h).
- Autonomous drone flights are possible under normal weather conditions.
- Drones have the capability to move between pre-established airstrips, where pilots can manually replace batteries or cargo for delivery, thus ensuring improved logistics and the range of package dispatch.

Among the drawbacks of the project, the following points should be noted:

- The cargo capacity of the drone is limited to 1.8 kg.

- Delivery to the destination occurs by dropping the package from an elevated position (20-35 m), and the pre-packed package descends slowly using a paper parachute. Consequently, the distance from the actual landing point to the anticipated one may vary up to 5 meters.
- The service primarily deals with the delivery of medicines or related medical items. The list of possible non-medical types of packages is limited to restaurant or grocery products, and the like.

It is also worth adding that the company has developed the next generation of drones that deliver cargo via a tether instead of deploying a parachute. However, this type of delivery significantly reduces the range and speed.

Users of the service can place orders on the company's website and monitor the delivery process through a dedicated mobile application.

2.2. Analysis of the latest research for drone delivery systems

The delivery by drones involves the management of a large number of Unmanned Aerial Vehicles (UAVs) simultaneously. The logistics challenges of such operations have been extensively discussed in [5], where a multi-physics model at the system level is presented for optimal control of multi-engine UAVs. This model can be utilized in the development and evaluation of control strategies. The authors demonstrate the capability of using this model for basic maneuvers and lay the groundwork for planning more complex maneuvers and complete missions. For instance, drones may employ different control strategies for achieving maximum energy efficiency under high and low battery levels.

The proposed system offers the following advantages:

- The multi-physics model enables a more comprehensive and accurate representation of the dynamics and interaction between different physical systems of multi-engine UAVs.
- An optimal control strategy is developed to minimize a cost function that considers time and energy consumption.
- The proposed methods are flexible and adaptable to various types of UAVs or other aerial systems, making them valuable for a wide range of applications in transportation, surveillance, mapping, etc.

During drone flight, a significant amount of computation is required to adjust the process and mission specifics of UAVs. The transfer of computational load from the drone to a cloud structure is discussed in [6]. The described framework features a client-server architecture, positioning the drone as a client and the cloud as a scalable server. Overall, it has potential applications in various fields requiring efficient drone management, especially in the delivery sector.

The proposed system has the following advantages:

- Scalability: The client-server architecture of the framework ensures effective communication between multiple drones and the cloud server, enabling real-time control of a large number of drones.
- Efficiency: By offloading certain tasks to the cloud server, the workload on individual drones is reduced, allowing them to operate more efficiently.

- Open source: The framework is open source, allowing developers to freely use and modify it.
- Versatility: The framework has potential applications in various industrial sectors.

A crucial aspect of a drone's mission during delivery is the route planning from the collection point to the delivery point. The optimization of the sequence of these processes is discussed in [7]. The delivery problem involves a group of couriers (drones) timely delivering orders to clients. The goal of the algorithm is to increase profit over a specific time interval and reduce the overall delivery time. The authors propose a Markov decision process model for the "courier" assignment task, using deep learning algorithms to address the problem in a dynamic environment. Successful implementation of this algorithm could significantly impact the delivery industry, enhancing its speed and increasing company profits.

An important task for optimizing the algorithm in a drone delivery system is to consider the drone's battery usage. Aiello et al. [8] presents a model of energy consumption for a similar urban logistics infrastructure. This methodology allows considering various factors affecting drone battery consumption, such as cargo weight, the size of the serviced urban area, population density, flight range, built-in battery capacity, etc. The model was developed to help researchers better understand the energy needs of delivery systems using UAVs and identify ways to optimize their performance.

2.3. Review of common approaches and algorithms for drone delivery route planning

The main stage in any type of cargo transportation is the process of route planning, for which there are currently numerous algorithms aimed at solving transportation problems efficiently and quickly. These algorithms are a crucial component of logistics and transportation infrastructure.

Such problems arise when it is necessary to determine the optimal delivery route from the point of origin to the destination, taking into account various external factors such as distance, cost, time constraints, and resources. As a result, this algorithmic process can become quite complex, especially when dealing with a large number of delivery points in complex urban conditions.

Navigating the UAV through the city is a tough task involving many safety preconditions, so an optimal way would be to deliver the packages to the delivery offices all around the city. Such an approach would allow to manually build the safe routes between many adjacent departments, thus creating a graph with nodes and branches of given cost (routes length), where we need to find a path to navigate.

Let's consider the most common approaches to solving such a problem.

2.3.1. Traveling Salesman Problem algorithm

One of the most common and straightforward methods for building a delivery route is the Traveling Salesman Problem (TSP) algorithm. It is based on a mathematical model that helps find the shortest path that connects all given pickup and delivery points. The TSP algorithm takes into account various factors, such as the distance between points, loading and unloading times, vehicle capacity constraints, and other limitations.

The main advantage of the Traveling Salesman Problem algorithm is its simplicity and ease of implementation. This algorithm uses a brute-force approach, where all possible combinations between points are considered. This makes it accessible for use in various fields and research and simplifies its integration with other parts of the software code.

There are many variations of the Traveling Salesman Problem-solving methods, such as the Monte Carlo method, the method of averaged coefficients, or the nearest neighbor method. The nearest neighbor method uses heuristic estimation in its calculations, significantly speeding up the search for the optimal route but not guaranteeing absolute optimality.

However, it is worth noting the disadvantages of this algorithm. Since the Traveling Salesman Problem algorithm, at each point, must choose the next point from those it has not yet visited, there are $(n-1)!$ routes for the asymmetric and $\frac{(n-1)!}{2}$ routes for the symmetric Traveling Salesman Problem. This means that the size of the search space depends exponentially on the number of points. For an average-sized problem, finding the optimal route can take an unacceptably long time, as most of it is spent on the enumeration of all possible combinations between points, which requires significant computational resources.

Another drawback of the Traveling Salesman Problem algorithm is that it provides only an approximate solution and does not guarantee finding the shortest path. Consequently, the accuracy of these calculations decreases proportionally as the problem size increases.

2.3.2. Dijkstra's algorithm

Dijkstra's algorithm is one of the most common algorithms for finding the optimal path in a graph and has broad applications in various fields, including telecommunications, transportation networks, routing, and logistics planning.

The working principle of the Dijkstra's algorithm involves iteratively updating the shortest distances from the initial node of the graph to all other nodes. During its operation, each vertex is examined, and the distance to adjacent vertices is calculated using the corresponding edges [9]. As a result of its work, the Dijkstra's algorithm not only determines the shortest distances from the initial node to all other nodes but also memorizes the corresponding routes.

The Dijkstra's algorithm is quite efficient and performs well at optimal scales. Its execution time depends on the number of vertices and edges in the graph, but with proper implementation, it has a time complexity of $O(n^2)$, where n is the number of vertices. Despite the fact that finding a route involves exploring all possible path variations, this feature can be considered an advantage to some extent because having data about all available routes guarantees the optimality of the found solution.

It is evident that as the number of vertices and edges in the input graph increases, exploring all possible variations will significantly slow down the process of finding the shortest path, rendering the algorithm unsuitable for use with such input data.

2.3.3. A* algorithm

The A* (A-star) algorithm is also aimed at finding a path in a graph and is an improved version of the Dijkstra's algorithm.

To achieve maximum efficiency with the A* algorithm, the heuristic function should be chosen according to the specific problem, as there is no one-size-fits-all solution. When tying the final path cost to the distance, more efficient heuristic functions such as the Euclidean distance or the Manhattan metric should become the preference.

One of the key advantages of the A* algorithm is its efficiency compared to the Dijkstra's algorithm. It uses a heuristic estimate (denoted as "h") to calculate the distance from the current node to the final destination. This heuristic helps the algorithm make decisions about which node is likely to lead to the shortest path. When the heuristic function is optimistic (i.e., it doesn't overestimate the distance), the A* algorithm guarantees finding the shortest path.

All of these factors make A* a popular choice and an efficient tool for route planning and optimization in various fields, including robotics, artificial intelligence development, and routing.

Another advantage of the A* algorithm is its ability to handle graphs of moderate size and relatively complex problems efficiently. While the computational complexity depends on the graph's size, A* demonstrates high efficiency with optimal implementation. It can quickly find the shortest path when using a heuristic function that provides spatial orientation information.

Therefore, this algorithm is faster and more optimized for larger tasks compared to the Dijkstra's algorithm or the Traveling Salesman algorithm since it doesn't require exploring all possible route combinations.

However, one significant drawback of the A* algorithm is its potential to get trapped in local maxima. This means that an incorrectly defined heuristic function or an insufficiently informative estimate of a particular distance can influence the algorithm to choose the wrong path, which consequently is not the shortest. This can be problematic, especially when solution accuracy is critical, such as in robotics or automated route planning.

Another issue with the A* algorithm is high memory usage. Since it keeps track of all visited nodes, the memory requirements for storing this information can significantly increase for large input graphs or complex-sized problems. This may necessitate size limitations on problems that can be effectively solved using this algorithm without compromising its performance.

2.3.4. Reinforced learning

The fourth algorithm, considered when choosing a method for constructing an optimal route, employs a reinforcement learning approach to build delivery routes. It is based on machine learning concepts and uses the learning process to make decisions regarding the selection of the shortest and most efficient routes [7].

One of the advantages of reinforcement learning algorithm is its ability to self-learn and adapt to a dynamic environment. It can interact with the environment, learn based on provided rewards, and refine its strategy over time. This allows the algorithm to effectively operate in dynamic and uncertain situations, where predefined rules may be insufficient or inefficient.

Another advantage of the reinforcement learning algorithm is its ability to optimally utilize resources [10]. It can find a balance between exploring new possibilities and exploiting existing knowledge, maintaining a trade-off between exploration and task execution. This makes it valuable for real-time decision-making and managing complex systems like drone delivery.

The main drawback of this algorithm is the need for a large amount of data and proper data preparation. Reinforcement learning algorithm requires an adequate quantity of high-quality

initial data for effective learning. Improper data preparation can lead to errors during the model training phase, as the algorithm is sensitive to noise.

2.3.5. Choosing the final algorithm

To choose the appropriate algorithm for solving the drone delivery route problem, we first identified the main criteria and requirements for the developed software. Let's examine the identified issues in detail:

- Execution speed: the selected algorithm should be highly efficient in terms of computation time, as this ensures reduced delays in drone management.
- Scalability: the limitations of the algorithm regarding its maximum computational capacity and the size of the problem it can handle should be taken into account, ensuring scalability.
- Implementation simplicity: due to the extensive work involved in creating the software for the drone delivery system, the chosen algorithm should be relatively easy to integrate with other software modules. Guided by these requirements and criteria, let's evaluate the suitability of the previously analyzed most common route optimization algorithms.

The traveling salesman algorithm aims to find the shortest path that passes through each node in the graph and returns to the starting node. While it guarantees finding the shortest path and is relatively simple to implement, its computational complexity increases rapidly with the number of delivery points. This, in turn, affects processing speed and scalability, making it potentially less suitable for large-scale problems.

Dijkstra's algorithm is a classic approach to finding the shortest path in a graph with non-negative edge weights. It works by layer-wise propagation from the starting node to the destination. Its effectiveness lies in its ability to find the shortest path to every node in the graph. However, as the number of nodes and edges grows, the exhaustive search of all possible route combinations slows down the search process, affecting both processing speed and scalability.

The A* algorithm combines ideas from Dijkstra's algorithm and heuristic methods. It uses estimates of distances to the destination to expedite the search process. It can find the shortest path when information about the graph's structure is available. A* is particularly useful in complex state space problems or situations with limited resources. However, its computational complexity depends directly on the efficiency of the heuristic estimate. Additionally, it may require significant memory resources during route computation, which correlates with the input problem's size.

Reinforcement learning is a different approach to solving route optimization problems. It's based on the idea of training a model through trial and error. An agent learns to make decisions based on rewards and penalties received during specific actions. This approach allows the agent to adapt to changing environmental conditions and seek optimal solutions. While reinforcement learning can be time and resource-intensive during the training phase, the computational demands are primarily associated with the training phase rather than the actual deployment.

Considering the outlined criteria and requirements, the choice of algorithm for solving the drone delivery route problem depends on the specific characteristics and constraints of the problem, the availability of domain-specific information, and the balance between computational

complexity and scalability. Each of the algorithms mentioned has its strengths and weaknesses, making them suitable for different scenarios. The selection should be driven by the specific needs and goals of the drone delivery system.

Based on the analysis of the mentioned algorithms, it can be argued that reinforcement learning is the most optimal solution for the drone delivery route problem. Its ability to self-learn and adapt to changing conditions makes it an ideal choice. Reinforcement learning enables the system to quickly and efficiently determine the best route, avoid obstacles, and optimize delivery. Considering the need for speed and accuracy, this algorithm will facilitate optimal delivery with minimal resource consumption.

3. System development

3.1. General system architecture

Within the research of the main algorithms for the drone delivery software, a necessary step is to identify its key structural elements, essential for the implementation and operation of the chosen algorithm, and the methods of communication between them. It has been determined that such software should consist of three main modules, which, during the actual implementation, form a client-server architecture. The architectural structure of the system is schematically depicted in figure 1.

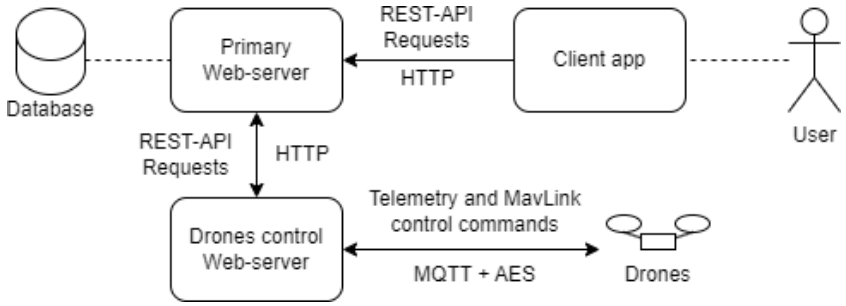


Figure 1: The architecture of the drone delivery system.

The client in this system is a user application, the purpose of which is to facilitate the user's interaction with the system.

Using the REST API interface, the application sends HTTP requests to a web server, which, in turn, processes them and performs necessary actions on the data, specifically basic CRUD operations (Create, Read, Update, Delete). This way, all client actions regarding interaction with the system, such as registration, creating or viewing lists of shipments, and their statuses, are handled.

A separate component of the system is the drone management module, which accumulates all the necessary methods and communication protocols. It also calculates optimal delivery routes. Utilizing the MQTT protocol [6], this part of the software creates a "Publisher-Subscriber" environment with the drones. This approach is commonly used in the Internet of Things [11, 12, 13] because it allows the server and hardware components to exchange messages freely

without the need for continuous monitoring of the system's status, as is required to receive updates through HTTP requests.

The data being transmitted consists of commands for drone control described using the "MavLink" protocol, which is a universal communication method with unmanned vehicles [14]. It enables both the sending of commands and receiving telemetry data from the drone, loading mission routes, and switching flight modes.

3.2. Hardware simulation

Due to limited testing capabilities caused by military actions in the territory of Ukraine, a crucial step is the selection of a sufficiently powerful and flexible technology for simulating the system's operation in a real environment. According to Chen et al. [15], one such technology is ArduPilot SITL, open-source autopilot software that allows the simulation of the control process for various types of unmanned vehicles, including drones. It provides access to a wide range of functionalities such as UAV mission planning, autonomous takeoff and landing, GPS waypoint navigation, and more. Thanks to ArduPilot, critical points of the system and the possibility of its further physical implementation can be easily assessed.

ArduPilot SITL also allows to monitor the position of a simulated quadcopter on an interactive map with satellite images, which comes very handy when looking on the actual navigation path in the cities.

Figure 2: ArduPilot SITL Map view with multiple simulated quadcopters [16].

3.3. Route building subprogram

The code blocks below show the final path finding code which was developed during the research work. It utilizes Q-Learning – a popular approach of reinforced learning that allows an agent to effectively learn efficient routes based on the given transportation costs in the graph. By tuning the hyperparameters the code was optimized to work reliably on any given set of waypoints.

The two main functions of a Q-Learning agent are the ones responsible for choosing an action and learning the consequences of executing a chosen action. Depending on a random choice and the current exploration probability (which decreases after each learning epoch) the agent will either choose a random action available from the current state, or utilize an action which brought the most reward during past iterations. On the each subsequent episode it will less likely explore the new moves and will instead exploit the collected route building data that is stored in his Q-table.

```
def choose_action(self, state):
    if np.random.uniform(0, 1) < self.exploration_prob:
        return np.random.choice(self.num_actions)  # Explore
    else:
        return np.argmax(self.q_table[state, :])  # Exploit
```

In order to remember the efficiency of all the action combinations, the agent is updating its Q-Table after taking each action by comparing the predicted outcome based on the past runs with the real reward obtained after the latest move.

```
def learn(self, state, action, reward, next_state):
    predict = self.q_table[state, action]
    target = reward + self.discount_factor *
        np.max(self.q_table[next_state, :])
    self.q_table[state, action] += self.learning_rate *
        (target - predict)
```

The overall learning process consists of moving across the graph and calculating the total cost of a route, repeating until a given amount of episodes (epochs) is not completed. By collecting the different rewards the agent is able to successfully learn the valid behavior that is leading him to maximum reward, thus finding the most optimal route.

```
# Make agent learn the graph for given episodes count
for episode in range(max_episodes):
    state = start
    total_reward = 0
    visited_nodes = []

    # While all necessary nodes are not visited
    while len(visited_nodes) != len(nodes_to_visit):
```

```
        # Choose a random move or exploit known data
        action = agent.choose_action(state)
        next_state = action

        # Add a negative reward for revisiting the same waypoint
        if action == state:
            reward = -100
        else:
            # Negative cost for shortest path
            reward = -graph[state, action]

            # Slightly increase the reward when agent visits a route he
            # had to visit and didn't visit yet
            if next_state in nodes_to_visit and
                not (next_state in visited_nodes):
                reward *= 0.001
                visited_nodes.append(next_state)
            elif next_state in visited_nodes:
               # Add a negative reward for revisiting the same waypoint
                reward -= 100

        agent.learn(state, action, reward, next_state)
        state = next_state
        total_reward += reward

    # Decay exploration probability
    agent.exploration_prob *= agent.exploration_decay
```

After the agent is done learning it is building the final path once again, which will eventually be the most efficient one, based on the pre-set hyperparameters and reward calculation logic.

```
# Graph array represents the costs for traveling from
# node A to B (graph[A][B])
graph = np.array([
    [0, 50, 20, 30, 40],
    [50, 0, 10, 30, 80],
    [20, 10, 0, 40, 10],
    [30, 30, 40, 0, 20],
    [40, 80, 10, 20, 0]
])

start = 0
nodes_to_visit = [4, 2] # Destination routes
visited_nodes = []
```

```
path = [start]
while len(visited_nodes) != len(nodes_to_visit):
    # Choose new actions until all nodes_to_visit are visited
    action = agent.choose_action(path[-1])
    if action in nodes_to_visit:
        visited_nodes.append(action)
    path.append(action)
print("Shortest Path:", path)
```

The example graph used in the code is assuming each waypoint is accessible from any other waypoint and the travel cost is the same when moving in both directions. In reality the departments graph could be dynamically generated by modifying the costs regarding the weather and wind directions, thus also optimizing the route built for the real world conditions. The agent itself could also utilize the drones battery level, maximum travel distance left, total cargo capacity and multi-package delivery optimizations in his reward system to even better improve the UAV path for maximum productivity.

4. Conclusion

The research allowed us to examine the overall aspects of drone delivery system creation. After analyzing the existing commercial systems and research in the area we were able to determine the crucial aspects of such systems and develop the necessary architecture. The key focus

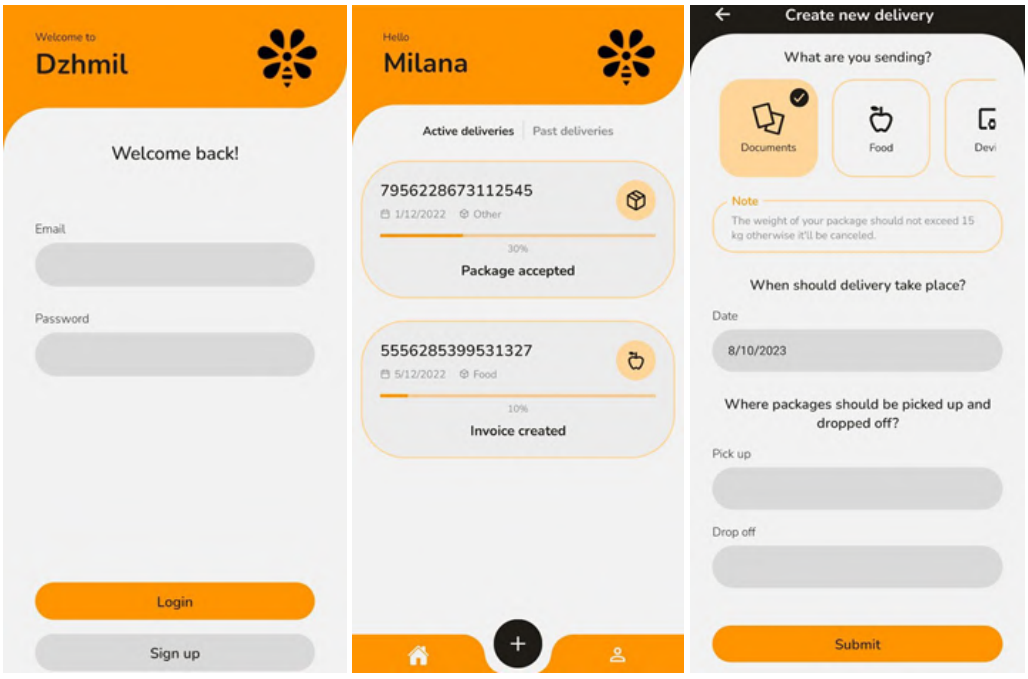


Figure 3: Examples of the user interface of the developed client application.

was given to the route planning algorithm that should create a valid path between start and destination location through a given set on departments (waypoints).

A backend and mobile applications were developed to be used by both regular users and delivery managers. The developed mobile app screenshots are displayed in figure 3.

In order to properly test all the system aspects, the drone flights were simulated in the ArduPilot SITL environment. This will also allow to directly apply the developed code to the UAVs running ArduPilot flight controller firmware.

Speaking of further development, the reinforced learning agent used for path finding can be improved by including different aircraft sensor data and environment conditions into the calculations. This will allow to embed such aspects as the payload weight, battery level or wind speed before the flight or even during the flight itself to better navigate the drone through the area.

The developed system had shown itself as a well working prototype that is easy to adapt and scale according to the desired conditions and requirements.

Acknowledgments

We acknowledge the contribution of ChatGPT to the refinement of this paper. ChatGPT's assistance in language enhancement and phrase generation significantly contributed to the quality of the final manuscript. It's imperative to highlight that the responsibility for reviewing and aligning the generated content with the narrative of our manuscript solely rests with the authors. We ensured that all content generated by AI tools, particularly regarding well-known concepts or definitions, underwent meticulous scrutiny to verify accuracy and relevance. Proper references to the original content were included to maintain academic integrity and acknowledge the intellectual contributions of others.

References

[1] A. R. Petrosian, R. V. Petrosyan, I. A. Pilkevych, M. S. Graf, Efficient model of PID controller of unmanned aerial vehicle, Journal of Edge Computing 2 (2023) 104–124. doi:10.55056/jec.593.

[2] Amazon Prime Air prepares for drone deliveries, 2022. URL: https://www.aboutamazon.com/news/transportation/amazon-prime-air-prepares-for-drone-deliveries.

[3] Starship Technologies: Autonomous robot delivery, 2023. URL: https://www.starship.xyz.

[4] Zipline Instant Delivery & Logistics, 2023. URL: https://www.flyzipline.com.

[5] N. Michel, Z. Kong, X. Lin, Optimal Control of a Multirotor Unmanned Aerial Vehicle Based on a Multiphysical Model, volume Volume 2: Intelligent Transportation/Vehicles; Manufacturing; Mechatronics; Engine/After-Treatment Systems; Soft Actuators/Manipulators; Modeling/Validation; Motion/Vibration Control Applications; Multi-Agent/Networked Systems; Path Planning/Motion Control; Renewable/Smart Energy Systems; Security/Privacy of Cyber-Physical Systems; Sensors/Actuators; Tracking Control Systems; Unmanned Ground/Aerial Vehicles; Vehicle Dynamics, Estimation, Control; Vibration/Control Sys-

tems; Vibrations of *Dynamic Systems and Control Conference*, 2020, p. V002T36A004. doi:10.1115/DSCC2020-3239.

[6] G. Mehrooz, E. Ebeid, P. Schneider-Kamp, System Design of an Open-Source Cloud-Based Framework for Internet of Drones Application, in: 2019 22nd Euromicro Conference on Digital System Design (DSD), 2019, pp. 572–579. doi:10.1109/DSD.2019.00087.

[7] H. Jahanshahi, A. Bozanta, M. Cevik, E. M. Kavuk, A. Tosun, S. B. Sonuc, B. Kosucu, A. Başar, A deep reinforcement learning approach for the meal delivery problem, Knowledge-Based Systems 243 (2022) 108489. doi:10.1016/j.knosys.2022.108489.

[8] G. Aiello, R. Inguanta, G. D'Angelo, M. Venticinque, Energy Consumption Model of Aerial Urban Logistic Infrastructures, Energies 14 (2021) 5998. doi:10.3390/en14185998.

[9] H. Huang, A. V. Savkin, C. Huang, Drone Routing in a Time-Dependent Network: Toward Low-Cost and Large-Range Parcel Delivery, IEEE Transactions on Industrial Informatics 17 (2021) 1526–1534. doi:10.1109/TII.2020.3012162.

[10] T. Lorido-Botran, M. K. Bhatti, ImpalaE: Towards an optimal policy for efficient resource management at the edge, Journal of Edge Computing 1 (2022) 43–54. doi:10.55056/jec.572.

[11] N. M. Lobanchykova, I. A. Pilkevych, O. Korchenko, Analysis and protection of iot systems: Edge computing and decentralized decision-making, Journal of Edge Computing 1 (2022) 55–67. doi:10.55056/jec.573.

[12] O. V. Klochko, V. M. Fedorets, M. V. Mazur, Y. P. Liulko, An IoT system based on open APIs and geolocation for human health data analysis, CTE Workshop Proceedings 10 (2023) 399–413. doi:10.55056/cte.567.

[13] Y. B. Shapovalov, Z. I. Bilyk, S. A. Usenko, V. B. Shapovalov, K. H. Postova, S. O. Zhadan, P. D. Antonenko, Harnessing personal smart tools for enhanced STEM education: exploring IoT integration, Educational Technology Quarterly 2023 (2023) 210–232. doi:10.55056/etq.604.

[14] A. Sharma, P. Vanjani, N. Paliwal, C. M. Basnayaka, D. N. K. Jayakody, H.-C. Wang, P. Muthuchidambaranathan, Communication and networking technologies for UAVs: A survey, Journal of Network and Computer Applications 168 (2020) 102739. doi:10.1016/j.jnca.2020.102739.

[15] W. Chen, Y. Dong, Z. Duan, DPM: Towards Accurate Drone Position Manipulation, IEEE Transactions on Dependable and Secure Computing 20 (2023) 813–826. doi:10.1109/TDSC.2022.3144319.

[16] Multiple Vehicles with MAVProxy, 2023. URL: https://ardupilot.org/mavproxy/docs/getting_started/multi.html.

Implementing E2E tests with Cypress and Page Object Model: evolution of approaches

Inessa V. Krasnokutska, Oleksandr S. Krasnokutskyi

Yuriy Fedkovych Chernivtsi National University, 2 Kotsiubynskoho Str., Chernivtsi, 58002, Ukraine

Abstract

This article shows eight approaches how to construct Cypress tests using POM. The connections between them are stressed as their evolution while writing code developing E2E tests. The authors highlight advantages and disadvantages of the approaches and offer the solution of problems. This article can be used both as a combined overview of different approaches and as a manual for those who are struggling to write tests with Cypress in a better way.

Keywords

Cypress JS, Page Object Model, POM design pattern, automation testing, E2E tests

1. Introduction

Suppose we have a problem to cover the functionality of certain website with automation tests. For instance we can consider the website https://www.saucedemo.com/v1/index.html without loss of generality. The website is devoted to illustrating of different test cases that occur during unsuccessful login process. It contains the form with username and password fields to get the access to the next page, and in case of unsuccessful log in, the error message with corresponding text is shown. This website is convenient to be used as a model example for all typical flows of users behaviour that could be covered by tests with a minimum amount of effort.

Successful login without any doubt as a positive test case means inputting correct username and password. Negative tests occur when a user is left on the same page and an appropriate error message is shown. It can happen when a user enters an empty or wrong username or password. Different tests can be considered to cover those cases. However, taking into account that most of developed tests are quite similar and optimising the cases authors demonstrate only valuable and significant of them in this article. More test cases can be found in the project repository at https://github.com/InaKrasnokutska/CypressPOM. To run the solution you just need to pull and run **npm i** to setup dependencies, and later **npm run cy:open** to see how it works.

Cypress is chosen to demonstrate how to create tests in different manners. All pros and cons of each approach are analysed and highlighted. There is a lack of good resources on the Internet

CS&SE@SW 2023: 6th Workshop for Young Scientists in Computer Science & Software Engineering, February 2, 2024, Kryvyi Rih, Ukraine

i.krasnokutska@chnu.edu.ua (I. V. Krasnokutska); okrasn.mail@gmail.com (O. S. Krasnokutskyi)

https://amit.chnu.edu.ua/pro-kafedru/personalii/krasnokutska-inessa-volodymyrivna (I. V. Krasnokutska)

0000-0002-7034-7291 (I. V. Krasnokutska)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

CEUR-WS.org/Vol-3662/paper24.pdf

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

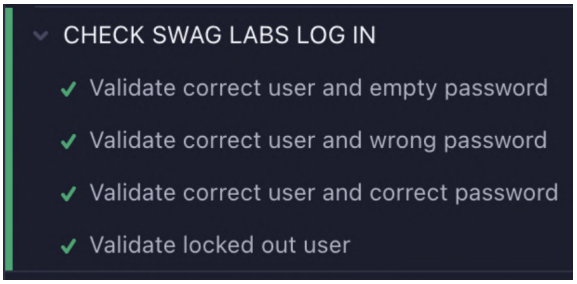


Figure 1: Test cases you can find at https://github.com/InaKrasnokutska/CypressPOM and run in Cypress.

about usage of the Page Object design pattern in Cypress Cypress [1], and each resource shows only one point of view and it is hard to understand the correlation between them.

2. Main results

All possible variations of implementing Page Object Model (POM) are constructed in this article. It happens when a locked out user enters his/her username and password and receives an appropriate error message.

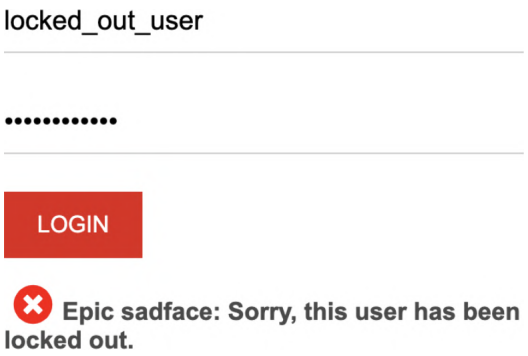


Figure 2: Error message the locked out user tries to log in.

Other test cases can be found at https://github.com/InaKrasnokutska/CypressPOM. In the next subsections we will discuss 9 approaches to organise selected test case. For convenience and quick search of files with tests (specs) in the repository they have names according to subsection numbers in this article where they are discussed.

2.1. Tests without POM

Before diving into the Page Object design pattern, let us create a Cypress test without it.

According to this decision the code appears very strict forward because in the spec file we need to prescribe the interaction with each DOM element of the page. This interaction is done

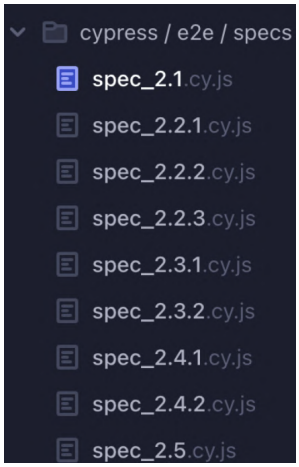


Figure 3: Spec's names corresponding to subsections where they were discussed.

with the help of the function **cy.get()** that locates an element with the corresponding selector of the element passed into it as the parameter.

Next we emulate the user's action by typing username and password and clicking the button (**.type()** and **.click()** functions). At the end of the scenario we validate the error message using **.should()** function for assertion.

Listing 1: File **spec_2.1.cy.js**

```
describe('CHECK SWAG LABS LOG IN WITHOUT POM 2.1', () => {
 it('Validate locked out user', () => {
   cy.get('#user-name').type('locked_out_user');
   cy.get('#password').type('secret_sauce');
   cy.get('#login-button').click();
   cy.get('[data-test="error"]').should('be.visible')
     .and('contain.text',
       'Epic sadface: Sorry, this user has been locked out.');
 });
});
```

This approach lets us to obtain expected result of testing but in the same time it does not fit list of quality criteria [2] because of

- complexity of maintenance and updating,
- level of duplication increase dependent on covered scenarios growth,
- violation of principles: KISS, DRY, SOLID etc.

2.2. Tests with POM using selectors for elements

For improving the code above let us use the Page Object design pattern and update the code according to style used in Selenium/Java [3].

2.2.1. POM using selectors

First of all, we describe the class that contains selectors of elements in variables. Then, implementing functions and locators we use selectors predefined in variables. Here locators are implemented with **cy.get()** help and are used to perform actions on page.

Listing 2: File **login_2.2.1.page.js**

```
class MainPage {
  usernameInputSelector = '#user-name';
  passwordInputSelector = '#password';
  loginButtonSelector   = '#login-button';
  errorMessageSelector  = '[data-test="error"]';

  typeUsernameInput(username)
    { cy.get(this.usernameInputSelector).type(username); }

  typePasswordInput(password)
    { cy.get(this.passwordInputSelector).type(password); }

  clickLoginButton()
    { cy.get(this.loginButtonSelector).click(); }

  checkErrorMessage(message)
    { cy.get(this.errorMessageSelector).should('be.visible')
        .and('contain.text', message); }
}
module.exports = new LoginPage();
```

As a result spec file will take the following form.

Listing 3: File **spec_2.2.1.cy.js**

```
import loginPage  from "../pages/login_2.2.1.page";

describe('CHECK SWAG LABS LOG IN WITH POM 2.2.1', () => {
 it('Validate locked out user', () => {
   loginPage.typeUsernameInput('locked_out_user');
   loginPage.typePasswordInput('secret_sauce');
   loginPage.clickLoginButton();
   loginPage.checkErrorMessage('Epic sadface: " +
     'Sorry, this user has been locked out.');
 });
```

Considering pros and cons of this approach, we can say that it isolates the logics of selecting elements apart from processing scenarios. This isolation improves disadvantages illuminated in the previous approach. At the same time this approach has disadvantage because assertions are performed by class methods, and it is better to leave assertions in the spec file.

2.2.2. POM using getter for error message

Thus, the next step of optimization is extension of class functionality for providing a possibility to make assertions within the spec file.

With this aim, let us define a function in class that will give access to the element, a kind of getter for error messages.

Listing 4: File **login_2.2.2.page.js**

```
getErrorMessage()
  { return cy.get(this.errorMessageSelector); }
```

Listing 5: File **spec_2.2.2.cy.js**

```
import loginPage  from "../pages/login_2.2.2.page";

describe('CHECK SWAG LABS LOG IN WITH POM 2.2.2', () => {
 it('Validate locked out user', () => {
   loginPage.typeUsernameInput('locked_out_user');
   loginPage.typePasswordInput('secret_sauce');
   loginPage.clickLoginButton();
   loginPage.getErrorMessage().should('be.visible')
     .and('contain.text',
       'Epic sadface: Sorry, this user has been locked out.');
 });
```

As a result of optimisation we obtained improvement of code. However, at the same time, the code is still not uniformed, as one element on the page has accessor function, and others do not have. Code can be understood better when we suppose that the necessity to call those elements directly in a spec file exists. For instance, when we need to verify that in case of unsuccessful username input the element changes its style (red color etc.).

2.2.3. POM using getters for all elements

Let us eliminate the shortcomings of the previous version and define functions for access to each page element.

Listing 6: File **login_2.2.3.page.js**

```
class MainPage {
  usernameInputSelector = '#user-name';
  passwordInputSelector = '#password';
  loginButtonSelector   = '#login-button';
  errorMessageSelector  = '[data-test="error"]';

  getUsernameInput(){ return cy.get(this.usernameInputSelector);}
  getPasswordInput(){ return cy.get(this.passwordInputSelector);}
  getLoginButton()  { return cy.get(this.loginButtonSelector);}
```

```
  getErrorMessage() { return cy.get(this.errorMessageSelector);}
  typeUsernameInput(username)
    { this.getUsernameInput().type(username); }
  typePasswordInput(password)
    { this.getPasswordInput().type(password); }
  clickLoginButton()
    { this.getLoginButton().click(); }
}
module.exports = new LoginPage();
```

Tests in **spec_2.2.3.cy.js** file are the same as in **spec_2.2.2.cy.js** file. At the same time, the new version contains more functionality for further extension and can cover bigger amount of user behavior scenarios with slight updates.

Evaluating the code, we observe that it currently seems overloaded with functions. That happens because we keep separately selectors of each element and create distinct functions to access each element that with help of selectors return us the locators of elements.

2.3. Tests with POM using locators for elements

One of the methods to simplify and optimise the code is the union of selectors and locators. It can be done because out of class only wrapped elements are used to perform actions on them. That is why there is no sense to keep distinct properties only for selectors.

2.3.1. POM using named functions as locators

Let us delete from the previous variant all the variables for selectors and strictly pass selectors as arguments in **cy.get()**. Each locator is obtained as a method that returns the element.

Listing 7: Part of file **login_2.3.1.page.js**

```
  getUsernameInput(){ return cy.get('#user-name');}
  getPasswordInput(){ return cy.get('#password');}
  getLoginButton   (){ return cy.get('#login-button');}
  getErrorMessage  (){ return cy.get('[data-test="error"]');}
```

We obtain 4 getter functions and focus on the necessary functionality. Then when we work with functions (**typeUsernameInput()**, **typePasswordInput()**, **clickLoginButton()**), we use the predefined locators of this class.

As a consequence the tests in **spec_2.3.1.cy.js** look the same as ones in **spec_2.2.2.cy.js** and work as well.

2.3.2. POM with anonymous functions as locators

Another way to define methods is the usage of anonymous functions. Function expressions are not hoisted, which means it is created only when the execution flow reaches it and can be used from that moment onwards.

Listing 8: Part of file **login_2.3.2.page.js**

```
getUsernameInput=function(){ return cy.get('#user-name');}
getPasswordInput=function(){ return cy.get('#password');}
getLoginButton  =function(){ return cy.get('#login-button');}
getErrorMessage =function()
                    { return cy.get('[data-test="error"]');}
```

The same as in the previous case, tests do not have changes (tests in **spec_2.2.3.cy.js** are the same as in **spec_2.2.2.cy.js**).

2.4. Tests using POM using locators in elements object

The usage of previous approach does not make semantic separation between different parts of the Page Object file, that is why we can perform aggregation and place locators into the **elements** object thereby dividing the functions into two groups: functions devoted to access elements and functions devoted to actions.

2.4.1. POM using locators as functions in elements object

Let us use getters names as the keys and locators functions as the values in **elements**.

Listing 9: File **login_2.4.1.page.js**

```
class MainPage {
  elements = {
    getUsernameInput: function(){ return cy.get('#user-name');},
    getPasswordInput: function(){ return cy.get('#password');},
    getLoginButton  : function()
                        { return cy.get('#login-button');},
    getErrorMessage : function()
                        { return cy.get('[data-test="error"]');}
  }
  typeUsernameInput(username)
    { this.elements.getUsernameInput().type(username); }
   typePasswordInput(password)
     { this.elements.getPasswordInput().type(password); }
   clickLoginButton()
     { this.elements.getLoginButton().click(); }
}
module.exports = new LoginPage();
```

According to this approach we refer firstly to the **elements** object and then to corresponding function every time we need to refer to the locators functions in code of action functions. Similar style we need to keep in the spec code also.

Listing 10: File **spec_2.4.1.cy.js**

```
import loginPage  from "../pages/login_2.4.1.page";
```

```
describe('CHECK SWAG LABS LOG IN WITH POM 2.4.1', () => {
 it('Validate locked out user', () => {
   loginPage.typeUsernameInput('locked_out_user');
   loginPage.typePasswordInput('secret_sauce');
   loginPage.clickLoginButton();
   loginPage.elements.getErrorMessage().should('be.visible')
     .and('contain.text',
       'Epic sadface: Sorry, this user has been locked out.');
 });
});
```

This approach can be found on some legacy projects or on projects not supporting modern stack.

2.4.2. POM using arrow functions as locators in elements object

Analyzing a block of elements we can note that it contains a long construction **function() { return ...}** and the idea to shorten or optimize it appears. It can be done if the project uses standard **JavaScript EcmaScript 6** or higher. Then, using arrow functions syntax we can obtain a more concise look of code.

Listing 11: Part of file **login_2.4.2.page.js**

```
 elements = {
   getUsernameInput: () => cy.get('#user-name'),
   getPasswordInput: () => cy.get('#password'),
   getLoginButton  : () => cy.get('#login-button'),
   getErrorMessage : () => cy.get('[data-test="error"]')
}
```

Due to point localization of the file, no updates needed for tests in **spec_2.4.2.cy.js** (they look the same as in **spec_2.4.1.cy.js**).

Investigating available articles on this topic we found out that most of the resources on the Internet recommend writing code with POM as it is shown in this subsection. However, in most cases they use incorrect names for functions in the **elements** object. The emphasis is on the subject and not on performing an action over the subject (verbs **get**, **take** etc. are omitted), for instance **usernameInput()**, **loginButton()**. When we need to refer to the part of elements in the tests definitely we want to name the elements as a noun, but round brackets for function call cannot be omitted. In that way code looks incomprehensible and weird violating name convention (**loginPage.elements.errorMessage().should('be.visible')**).

On the one hand this kind of names has benefits because they are used to denote locators of elements. But on the other hand it has the drawback: this part of code is performing some action and first of all it is a function and it should respect the name conventions for functions.

2.5. Tests with POM using assessor properties

The drawback described above can be avoided by using assessor properties. They are property getters, new types of properties (along with the regular data properties). They are essentially functions that execute on getting value, but look like regular properties to an external code. “getter” methods are accessor properties . In an object literal they are denoted by **get**. Descriptor for accessor properties has **get** – a function without arguments, that works when a property is read.

Listing 12: File **login_2.5.page.js**

```
class MainPage {
  get usernameInput()   { return cy.get('#user-name');}
  get passwordInput()   { return cy.get('#password');}
  get loginButton()     { return cy.get('#login-button');}
  get errorMessage()    { return cy.get('[data-test="error"]');}

  typeUsernameInput(username)
    { this.usernameInput.type(username); }
  typePasswordInput(password)
    { this.passwordInput.type(password); }
  clickLoginButton()
    { this.loginButton.click(); }
}
module.exports = new LoginPage();
```

Listing 13: File **spec_2.5.cy.js**

```
import loginPage  from "../pages/login_2.5.page";
describe('CHECK SWAG LABS LOG IN WITH POM 2.5', () => {
 it('Validate locked out user', () => {
   loginPage.typeUsernameInput('locked_out_user');
   loginPage.typePasswordInput('secret_sauce');
   loginPage.clickLoginButton();
   loginPage.errorMessage.should('be.visible')
     .and('contain.text',
       'Epic sadface: Sorry, this user has been locked out.');
 });
```

The getter works when **errorMessage** is read in spec. From the outside, an accessor property looks like a regular one. That’s the idea of accessor properties. We don’t call **errorMessage** as a function, we read it normally: the getter runs behind the scenes.

A minor drawback of this approach is that sometimes it is necessary to pass a parameter to the locator functions (for example referring to the **i**-th row of the table), then we need to combine the use of ordinary functions and assessor properties.

3. Conclusions

This article shows eight approaches how to construct Cypress tests using POM. The connections between them are stressed as their evolution while writing code developing E2E tests. Project using approaches highlighted in first subsections still can be found on the Internet while studying how to write tests with Cypress JS. The authors highlight advantages and disadvantages of the approaches and offer the solution of problems. This article can be used both as a combined overview of different approaches and as a manual for those who are struggling to write tests with Cypress in a better way. For further investigation other patterns and solutions like Cypress commands, aliases, application actions, could be discovered.

References

[1] J. T. Othayoth, S. Anuar, Modern web automation with cypress.IO, Open International Journal of Informatics 10 (2022) 182–196. URL: https://oiji.utm.my/index.php/oiji/article/view/229.

[2] M. Leotta, M. Biagiola, F. Ricca, M. Ceccato, P. Tonella, A Family of Experiments to Assess the Impact of Page Object Pattern in Web Test Suite Development, in: 2020 IEEE 13th International Conference on Software Testing, Validation and Verification (ICST), 2020, pp. 263–273. doi:10.1109/ICST46399.2020.00035.

[3] E. Vila, G. Novakova, D. Todorova, Automation Testing Framework for Web Applications with Selenium WebDriver: Opportunities and Threats, in: Proceedings of the International Conference on Advances in Image Processing, ICAIP '17, Association for Computing Machinery, New York, NY, USA, 2017, p. 144–150. doi:10.1145/3133264.3133300.

Design and development of a game application for learning Python

Vasyl P. Oleksiuk[1,2], Dmytro V. Verbovetskyi[2] and Ivan A. Hrytsai[1]

[1]Ternopil Volodymyr Hnatiuk National Pedagogical University, 2 Maxyma Kryvonosa Str., Ternopil, 46027, Ukraine
[2]Institute for Digitalisation of Education of the NAES of Ukraine, 9 M. Berlynskoho Str., Kyiv, 04060, Ukraine

Abstract

This article explores the design of a Python learning game application, presenting outcomes in meeting its objectives. The study involves an analysis of various experiences in game-based learning, leading to the substantiation of requirements and the development of a model for the application. The authors identified the basic requirements for a game application, such as relevance to educational objectives, incorporation of game mechanics, the ability to write and debug code in the application, individualized passing of the game, and the ability to report and rate learning outcomes of students. The authors identified the basic requirements for a game application, such as relevance to educational objectives, incorporation of game mechanics, the ability to write and debug code in the application, individualized passing of the game, and the ability to report and rate learning outcomes of students. Unity3D emerges as the chosen engine following comparative analysis, considering graphic development, interaction, and user-friendliness. The authors of the article described their own experience in developing the PythonLeaner game. The designing process of this game emphasises the incorporation of game elements with both test and open-ended tasks to enhance programming education. The paper contains some fragments of the game workspace and code of a gaming application. The PythonLeaner's game limitations of the game and further directions for overcoming them for improving the application are indicated.

Keywords

game-based learning, game engine, Unity, Unreal Engine, CryEngine, game, game development,

CEUR-WS.org/Vol-3662/paper25.pdf

1. The problem statement

Today, in the digital age, artificial intelligence game-based educational tools are widely used. Teachers use them as desktop and mobile applications, cloud services, LMS modules, etc. [1, 2]

The design of game applications to teach many science subjects especially computer science offers significant benefits, making it highly relevant for several reasons. First of all by utilising game applications for educational purposes, instructors are able to capture and sustain pupils' interest and motivation, simplifying the process of immersion into the subject matter [3]. Furthermore, game-based learning promotes active engagement, providing an effective method to

CS&SE@SW 2023: 6th Workshop for Young Scientists in Computer Science & Software Engineering, February 2, 2024, Kryvyi Rih, Ukraine

oleksyuk@fizmat.tnpu.edu.ua (V. P. Oleksiuk); verbovetskyj.dv@gmail.com (D. V. Verbovetskyi); grytsaj_ia@fizmat.tnpu.edu.ua (I. A. Hrytsai)
https://tnpu.edu.ua/faculty/fizmat/oleksyuk-vasil-petrovich.php (V. P. Oleksiuk)
0000-0003-2206-8447 (V. P. Oleksiuk); 0000-0002-4716-9968 (D. V. Verbovetskyi); 0000-0001-6261-8429 (I. A. Hrytsai)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

encourage student participation [4]. Students should participate in the game, solve problems and make decisions to encourage hands-on, experiential learning, which is a key aspect of mastering programming concepts. Games commonly require problem-solving, logic, algorithms, and critical thinking [5]. They are fundamental skills in science learning and programming. Game applications have the ability to simulate real-life scenarios and challenges that programmers are likely to face in their future careers. This feature helps students to connect theoretical knowledge with practical application. Games offer prompt feedback, allowing students to observe the consequences of their choices and make necessary modifications. This iterative learning process is in line with the problem-solving nature of programming. Many games feature a gradual difficulty curve, enabling students to begin with simple tasks and gradually advance to more intricate challenges. This approach mirrors the incremental learning method often used in programming [6].

Multiplayer modes of modern games can facilitate the development of teamwork and communication skills, which are vital in software development and real-world projects. Competitive games can inspire productive competition, motivating players to enhance their abilities [7]. Self-paced learning is also a key component of this approach. Game-based learning enables self-paced learning, enabling learners to progress through the content at their own speed, thus accommodating diverse learning styles and abilities. Some game-based platforms include coding elements, which allow learners to script or modify game behaviour. This hands-on exposure to coding strengthens programming concepts and syntax. Game-based learning can offer increased accessibility as it caters to a range of learning styles and can be tailored to suit different abilities and skill levels [8]. Moreover, it can provide better retention of information over the long-term. When students derive enjoyment from the learning experience, they tend to remember and employ what they have learned more effectively.

Although numerous ready-made tools are available on the internet [9], the issue of creating new instances remains relevant. So the aim of this research is to develop a model of a game application for learning Python. To achieve this, the following tasks need to be solved:

1. Analyse the experience of using game applications in teaching programming.
2. Describe the requirements for the application and design its model.
3. Select tools for application development.
4. Describe and analyse the key points of application development.

2. The model of the game application for learning Python

Based on extensive research by Priyaadharshini et al. [10], Kroustalli and Xinogalos [11], Papadakis et al. [12], Osadcha and Osadchyi [13], Yatsenyak et al. [14] it can be concluded that powerful learning game applications must meet specific, key requirements.

- The application should possess well-defined educational objectives, including imparting knowledge of programming concepts, problem-solving abilities, and coding languages, in accordance with established academic standards.
- The application should have elements of gamification. It is necessary to incorporate game mechanics, including challenges, rewards and progression systems, to enhance the

learning experience by making it engaging and enjoyable. Therefore, the game should be based on an engaging narrative or storyline that contextualises the learning content in the game.

- The application should be in a Hands-on style. For example offers chances for learners to write and debug code within the application. It is advisable to integrate coding challenges, puzzles, or issues that necessitate learners to put programming principles into practice.
- The game should be played individually by each student. To develop programming skills, the game should be played over several levels of increasing complexity. The tasks should comprehensively cover a range of programming concepts, from basic syntax to more advanced ideas, catering to the intended audience.
- Additionally, the learning application utilizes analytics and reporting to provide insightful measures of progress. Provide educators and learners with performance data and identify areas for improvement.
- It would be good if the app is available across multiple platforms (web, mobile, and desktop) to maximize the potential audience. In particular, the game must perform effectively on various screen sizes and devices.
- The application should permit users to access and learn from the application offline, allowing it to be utilized in regions with limited or no internet connectivity.

To develop the game model, the content, goal, objective, tasks and scheme of passing this game were developed. The aim of the game is to facilitate game-based learning [11, 15, 16] for first-year students taking the Python programming course. The PythonLeaner game we are developing should provide the student with several modes. They are implemented by menu items such as:

1. New game.
2. Continue.
3. Shop.
4. Exit.

The model of the game is shown in figure 1.

The item “New game” displays to the user a greeting. It introduces the player to the purpose of the game and introduces its rules. The next screen shows an assistant who will accompany the player throughout the game. The assistant has the appearance of a stylized snake and familiarizes the player with the game process in more detail. As can be seen from figure 1 the PythonLeaner game provides completing by users several levels (stages). There are 3 tasks in each level. After successful completion of each of the tasks points are awarded to the student. At any moment of the game, the player can open a “shop” purchase a computer upgrade and stylize its appearance.

If the player gives an incorrect answer during the tasks, the assistant will provide a hint that will help to get to the correct answer. The game process consists of theory and practice, the incentive element is that the player receives points after correct answers, which can be spent on the graphic design of the working environment.

During the game, the player has some options such as improving the appearance of his/her own computer. In particular, he/she can buy PC peripherals of a different appearance and can

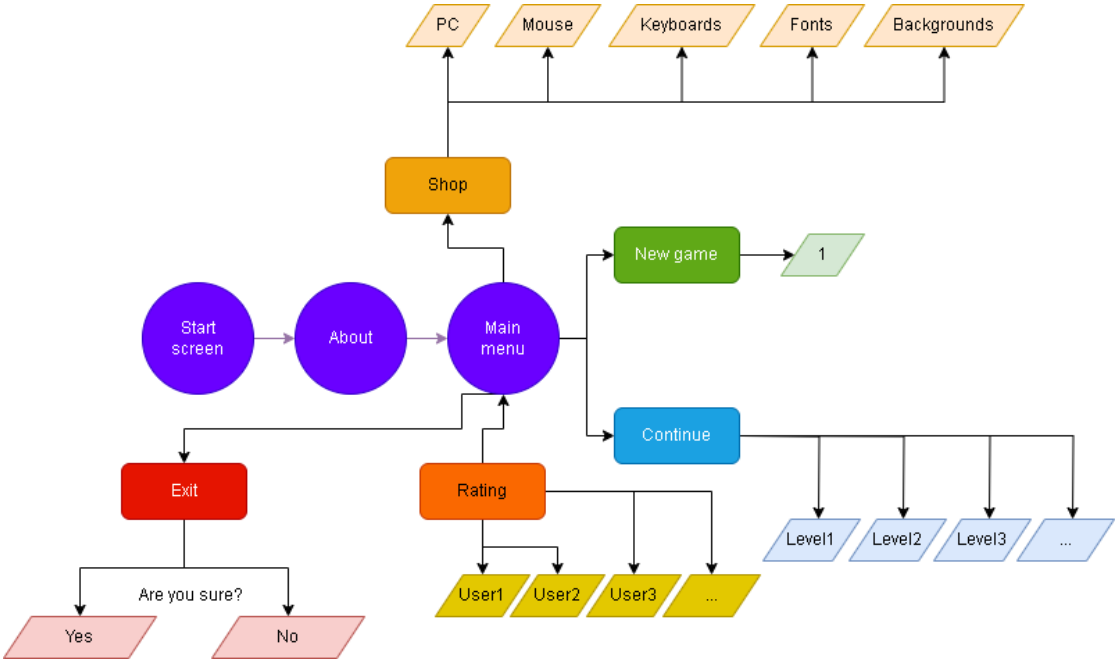


Figure 1: The model of the PythonLeaner game.

change the output devices (monitor, mouse, keyboard) the terminal font and the background of the working environment.

Moreover, it implemented an element of computer freezing. At the start levels, the user's computer freezes when processing the answer to a test question. The performance of the computer can be increased in the store by using the collected points. Finally, after completing all the levels, the player can get an infinite number of points and can switch between all the allowed interfaces. All player results are stored in the cloud. They are used to generate a ranking (figure 2), which is also available to players.

This approach makes it possible to implement a competitive approach to learning programming.

3. Programming environments for developing game applications

Each game is developed on a specific game engine and unified according to specific tasks. Action games are among the most popular [17]. This genre has a lot of subgenres such as platform games, fighting games, and, of course, shooting or shooters, which have not lost their popularity since the 1990s. This genre also includes arcade games, survival, sports games, simulators, and racing. The next most popular voluminous genre with a large number of subgenres is strategies. Strategies are economic, defensive towers, military, card. And the last of the most popular genres of games are role-playing games or RPGs [18]. These include such subgenres as educational, adventure, quests, puzzles, browser games, network text and network role-playing games. The game we are designing can be categorised as an educational quest.

Figure 2: The rating of PythonLeaner's gamers.

Let's consider in more detail some engines for creating games such as Unity, Unity3D, and CryEngine [19]. *Unity* environment makes it very convenient to use a 2D template to quickly start working on a project. This template is accessible via the Unity Hub and features default configurations, such as [20, 21].

1. A default scene that uses 2D mapping and contains the camera in orthographic mode with a single-colour sweep of the previous frame (not a Skybox, as is typical for 3D projects).
2. The editor is set to 2D mode by default, so new texture assets are imported as Sprites.
3. The global lighting mechanism is disabled.
4. A number of installed packages, including 2D Animation, 2D Pixel Perfect, 2D PSD Importer and 2D SpriteShape, as well as required dependencies.

Unity allows you to develop games for various platforms such as Windows, macOS, Linux, Android, iOS, Xbox and PlayStation consoles, virtual reality and others. This engine supports several programming languages, including C# and JavaScript. Engine has a large and active developer community, and there are plenty of resources such as forums, tutorials, and documentation to help you solve problems and learn new things. Unity provides a free version for small teams and independent developers. This promotes the spread of the engine and allows startups and individual developers to effectively enter the game development industry.

The platform is constantly being updated. Its developers adding new features and improvements. The regular update releases allow programmers to take advantage of the latest features. Unity provides powerful tools for working with graphics and animation, including realistic lighting effects, a particle system, and tools for modeling 3D scenes. Unity supports virtual reality and augmented reality well, allowing developers to create immersive games and applications for a variety of devices. The platform has become one of the most popular game engines. It is so due to its affordability, powerful tools, and extensive opportunities for developers. The advantages of this engine are:

- profitable solution from a financial point of view, since there is a free version of the program;

- easy mastering of the engine thanks to a wide selection of video lessons on the Internet;
- a fairly wide community between engine developers and users, so developers quickly correct errors and shortcomings in its operation.

A drawback of Unity is the amount of space that any game written using this engine takes up. For example, even a simple pixel computer game takes up quite a lot of computer RAM. The Cities Skylines, Hearthstone, Call of Duty and others well known games have been developed using Unity.

Unreal Engine is a powerful game engine used to develop video games and interactive virtual environments. Drivers are trying to make the most efficient use of computer architecture to optimize their driver's performance. Since 2015, Unreal Engine 4 has become free for game developers. Known for its powerful graphics and photorealistic rendering quality, this engine uses its own rendering engine that supports various lighting effects and shaders. Unreal Engine uses the C++ for game logic and development. Users can use either C++ or the built-in scripting language Blueprint. The second one is a visual programming language. The platform allows developers to create games for various platforms, including Windows, macOS, Linux, Android, iOS, Xbox and PlayStation consoles. There is extensive documentation and tutorials to help developers learn and improve their skills. Unreal Engine supports virtual and augmented reality well. It provides tools to create impressive VR games and applications [22]. Unreal Engine is free to use. Game developers only pay if their product is successful in the marketplace. The advantages of this engine are:

- a large community of developers and, as a result, a considerable number of video materials to illustrate the development process;
- item technical support at a fairly high level;
- debugged update mechanism;
- at least one new tool for work is added during each update;
- a large set of tools for creating animations.

However, there are also drawbacks to using Unreal Engine, including its unsuitability for developing simplistic games and the requirement to pay a five percent tax after the game becomes profitable.

CryEngine is a commercial game engine from Crytek. This engine offers a large set of tools for creating a real-time PC game. Planform contains many advanced graphics, physics and animation technologies, as well as many gameplay improvements. CryEngine is famous for its impressive graphics and detailing of large open worlds. It uses a powerful rendering engine and supports a wide range of impressive effects such as shadows and lighting. Platform supports virtual reality, allowing developers to create immersive VR projects and games [23]. Engine specializes in creating large, open and immersive worlds that allow players to freely explore the game. CryEngine supports game development for various platforms such as PC, Xbox and PlayStation consoles. Engine has a large developer community and documentation to help beginners and professionals use the engine. CryEngine has been used to create famous games such as Far Cry, Crysis, and the Robinson: The Journey series. The engine includes a powerful physics system that allows you to realistically simulate the movement and interaction of objects in the game. CryEngine is defined by its ability to create stunning open worlds and realistic

graphics. This makes it popular with developers looking to create the most immersive gameplay experiences. Game developers and researchers highlight the following advantages of this engine [24, 25]:

- some functions such as Flowgraph and Fmod provide to developers with excellent graphics;
- the platform provides simple mechanisms for creating artificial intelligence;
- this engine is more advantageous in terms of water and weather effects, which can make it suitable for games with realistic atmosphere;
- support for realistic modeling of destruction and object physics, which can be useful for creating interactive and dynamic game worlds.

Among the drawbacks, significant ones include enduring prolonged wait times for technical support with the free version, a community of users growing at a sluggish pace, and obsolete or absent documentation for numerous modules [26].

Table 1
Comparison of game engines Unity3D, Unreal Engine, CryEngine.

Criterion	Unity3D	Unreal Engine	CryEngine
Ease of learning	High	Average	High
Cost	Free	Free	License
Community and support	Big and active	Big and active	Present
Extensions and modding	Yes	Yes	Yes
Compatibility and ecosystem	Universal	Wide functionality	Universal
Multi-platform	Yes	Yes	Yes
Productivity and optimization	Depends on the developer	Depends on the developer	Capable of optimization
Multiplayer and network support	Yes	Yes	Yes
Artificial Intelligence	Yes	Yes	Yes

Based on the data from table 1, we have selected the Unity3D platform for developing the PythonLeaner game.

4. Analysis of key points of the game development

The development of the game took place in several stages, including the design of scenes (greetings, shop, selection of levels, etc.), filling of scenes, creation of scripts for transition between scenes, selection of levels, and creation of level scenes. When logging into the game, the user has the option to register and start a new game. Players who are already registered can log in with their login and password and resume the game from the level they have previously completed. User data, including login and password, as well as their rating, are stored using Firebase web hosting. Firebase is Google's mobile and web development platform that assists

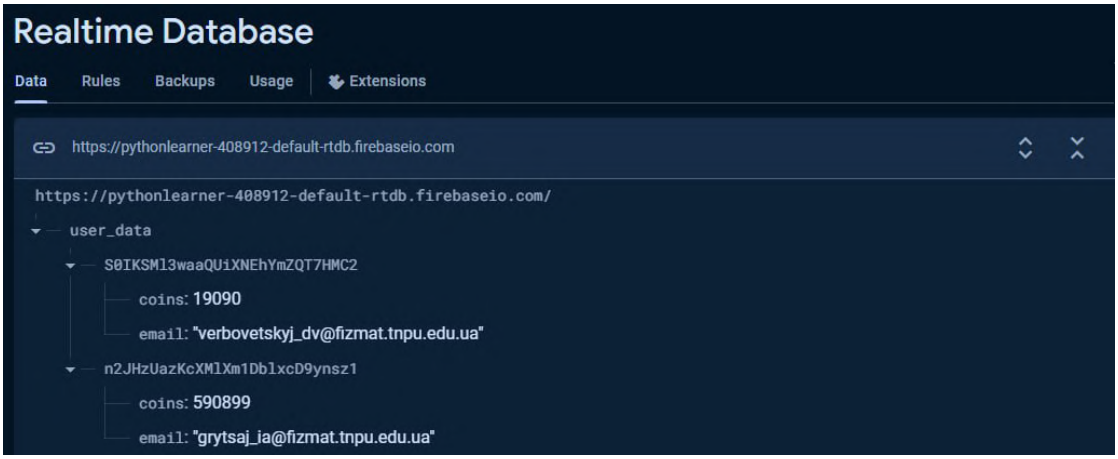


Figure 3: Rating saved in Firebase.

developers in building and scaling software products. The database for storing the rating required the installation of the SDK (FirebaseDatabase.unitypackage), as shown in the figure 3.

The game also has an option to choose the language of the interface. Users can switch between Ukrainian and English interface languages.

The IdleAnimation(), LikeAnimation(), DislikeAnimation() methods were used to implement the action of assistant (figure 4).

Figure 4: Reaction to the player's response.

These functions implement animation in three states of the assistant: "calm", "like", and "dislike". Each state changes the sprite. The Update() method is a main component in the code and is called with every frame. The method processes player keystrokes and updates the displayed text accordingly. This applies to deleting a character, wrapping a line, and other keys that may cause changes to the text. The following method, OnButtonClick(), is triggered when the player presses a button in the game. It compares the text entered by the player with the expected result. If the text is correct, the "like" animation is triggered. Otherwise, the "dislike" animation is displayed. The method performs changes in the game such as going to the next

scene.

To create the game, static variables are also used. One of them is coins (a static variable that stores the number of coins in the game and is available globally to all other objects and scripts). Others variables are pcPositions, pcSizes, promptPositions, promptSizes. These static lists of vectors and sizes define positions and sizes of objects in the game, such as computers. The selectedPCIndex is a static variable that stores the index of the selected computer in the list. The TextMeshProUGUI:mesh is a text field used to display the number of coins in the game.

An example of using the SetPCSizeAndPositionScript method is shown in Listing 1.

Listing 1: SetPCSizeAndPositionScript method

```
using System.Collections;
using System.Collections.Generic;
using UnityEngine;
public class SetPCSizeAndPositionScript : MonoBehaviour {
    void Start()  {
        RectTransform t = GetComponent<RectTransform>();
        t.sizeDelta = SingletonScript.pcSizes
                      [SingletonScript.selectedPCIndex];
        t.localPosition = SingletonScript.pcPositions
                          [SingletonScript.selectedPCIndex];
    }
    // Update is called once per frame
    void Update() {
    }
}
```

The project and gameplay have been depicted in figure 5. The user receives a default computer in the first level. The figure illustrates the default keyboard, mouse, font, terminal and workplace backgrounds, and an assistant (snake) located at the top right corner of the workplace. The player can request a hint from the assistant, and the snake's response depends on the student's answer. As you can see from the figure 5, the PythonLearner game has not only a desktop but also a mobile interface.

Tasks for the PyntonLeaner game are of several types such as tests (with one and several answers), entering short answers and writing Python's code. The essence of the game is that the player sees the screen of an old computer and an "improvised programming environment/terminal" is opened on it. During the first level, the player has an old computer a normal room background and a delay of 2.5 seconds when pressing the answer key, however, when completing the following levels, there is an opportunity to upgrade it by changing the appearance and reducing the delay when pressing the keys on the keyboard. The delay is implemented in such a way that when choosing each subsequent computer with a better design, its performance also increases. That is, each subsequent one from the computer by default has a delay of 0.5 seconds less than the previous one. The implementation of the key delay is described in listing 2. Each level consists of three types of tasks, including writing code. During the development of the project, it was decided to adapt the application so that new tasks could be created for

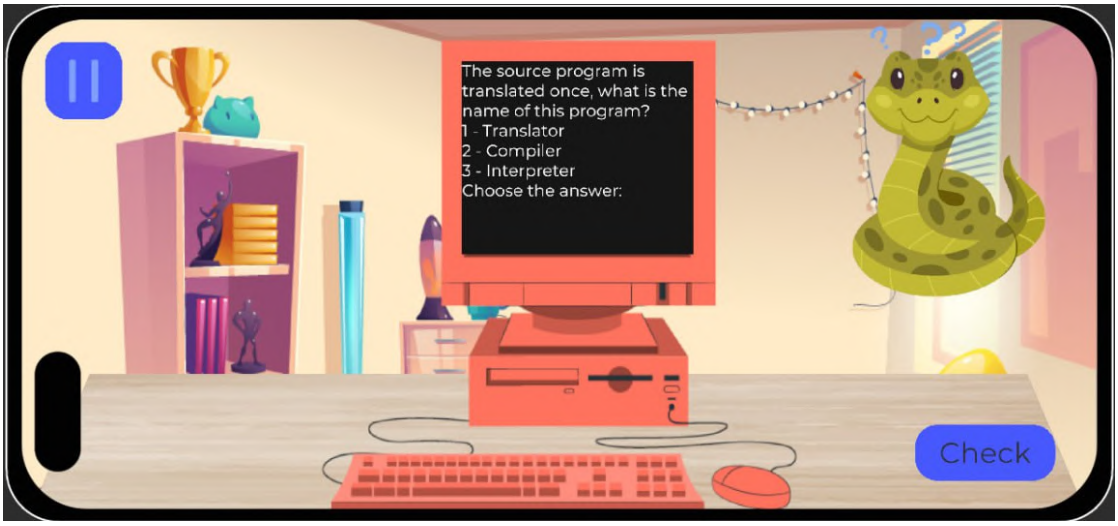


Figure 5: Gameplay of PythonLeaner.

learning other popular programming languages without changing the code. To check the code entered by the player, binding to each level of answers in the game engine was implemented, but this option is not very good, therefore, in the future of the game, we plan to implement an interpreter-based or artificial intelligence system in the game to automatically check the player's code. This innovative technology will allow us to effectively detect errors, optimize the work of players and ensure a high level of security in the game. The compiler or artificial intelligence will be responsible for analyzing the syntax, structure, and efficiency of the code, thus ensuring a higher quality of game software and reducing the possibility of vulnerabilities or anomalies.

Listing 2: Implementation of delay of input keys

```
public void OnButtonClick()      {
    Task.Delay(Convert.ToInt32
            (SingletonScript.pause * 1000)).Wait();
    if (text == result) {
        StopCoroutine(coroutine);
        StartCoroutine(LikeAnimation());
        SingletonScript.coins += 50;
    }
    else {
        StopCoroutine(coroutine);
        StartCoroutine(DislikeAnimation());
    }
}
```

The player learns the following from the assistant:

- the game consists of several levels, at each of them the player will be expected to have

theory and several practical tasks;

- after completing each level, the player receives a specific number of points. These points can be exchanged in the future for the graphic design of the game environment. The elements can be purchased at any time in the store (figure 6).

Figure 6: PythonLeaner Game Store.

After completing all the levels, the player gets an infinite number of points and can buy himself the best computer worth shown in figure 6. It costs 850 points and looks like a modern computer with a curved screen, backlit keyboard and cooling system. The player can choose the best background for the workplace after completing all levels (figure 7). This background

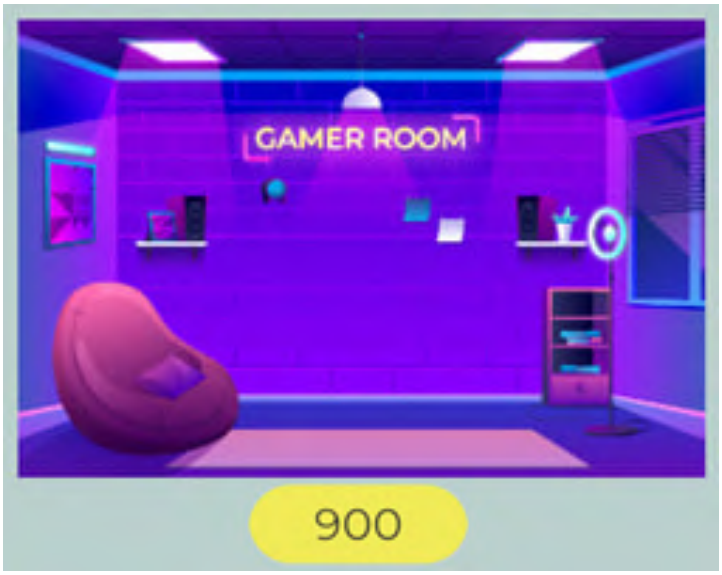


Figure 7: The best PythonLeaner desktop background.

is the most expensive of the ones offered and immerses the player in the real atmosphere of a programmer. There is no delay for entering characters from the keyboard during the passage of levels.

5. Conclusions

The results of this study show that its objectives were met. We analysed some experiences of creating and using game-based learning. On this basis, the requirements for a game application were substantiated and its model was developed. As a result of the comparative analysis, Unity3D was chosen as the main engine for the development of the PythonLeaner game. The chosen set of tools should take into account the needs of graphic development, interaction and ease of use. The author's experience in creating the game was systematised and the key points of its development were described. As a result of understanding the peculiarities of teaching programming, game elements with test and open-ended tasks were developed. It is important to emphasise the adaptability of the application. This implies the ability to adapt it to add tasks or learn other programming languages without editing the game code. Prospects for further research include the development of multiplayer, the integration of an interpreter to check the player's code, and the introduction of artificial intelligence. It is also planned to complete the game by creating a mobile application.

References

[1] O. V. Prokhorov, V. O. Lisovichenko, M. S. Mazorchuk, O. H. Kuzminska, Implementation of digital technology for student involvement based on a 3D quest game for career guidance and assessing students' digital competences, Educational Technology Quarterly 2022 (2022) 366–387. doi:10.55056/etq.430.

[2] E. Polat, Gamification implementation for educational purposes: a scoping review (2013-2018), Educational Technology Quarterly 2023 (2023) 367–400. doi:10.55056/etq.589.

[3] F. S. Breien, B. Wasson, Narrative categorization in digital game-based learning: Engagement, motivation & learning, British Journal of Educational Technology 52 (2021) 91–111. doi:10.1111/bjet.13004.

[4] A. I. A. Jabbar, P. Felicia, Gameplay Engagement and Learning in Game-Based Learning: A Systematic Review, Review of Educational Research 85 (2015) 740–779. doi:10.3102/0034654315577210.

[5] D. B. Clark, E. E. Tanner-Smith, S. S. Killingsworth, Digital Games, Design, and Learning: A Systematic Review and Meta-Analysis, Review of Educational Research 86 (2016) 79–122. doi:10.3102/0034654315582065.

[6] M. Priyaadharshini, N. Natha-Mayil, R. Dakshina, S. Sandhya, R. Bettina Shirley, Learning Analytics: Game-based Learning for Programming Course in Higher Education, Procedia Computer Science 172 (2020) 468–472. doi:10.1016/j.procs.2020.05.143.

[7] M.-T. Cheng, W.-Y. Huang, M.-E. Hsu, Does emotion matter? An investigation into the relationship between emotions and science learning outcomes in a game-based learning

environment, British Journal of Educational Technology 51 (2020) 2233–2251. doi:10.1111/bjet.12896.

[8] M. Barr, Student attitudes to games-based skills development: Learning from video games in higher education, Computers in Human Behavior 80 (2018) 283–294. doi:10.1016/j.chb.2017.11.030.

[9] J. Díaz, J. A. López, S. Sepúlveda, G. M. Ramírez Villegas, D. Ahumada, F. Moreira, Evaluating Aspects of Usability in Video Game-Based Programming Learning Platforms, Procedia Computer Science 181 (2021) 247–254. doi:https://doi.org/10.1016/j.procs.2021.01.141.

[10] M. Priyaadharshini, M. Natha, R. Dakshina, S. Sandhya, B. Shirley, Learning Analytics: Game-based Learning for Programming Course in Higher Education, Procedia Computer Science 172 (2020) 468–472. doi:10.1016/j.procs.2020.05.143.

[11] C. Kroustalli, S. Xinogalos, Studying the effects of teaching programming to lower secondary school students with a serious game: a case study with Python and CodeCombat, Education and Information Technologies 26 (2021) 6069–6095. doi:10.1007/s10639-021-10596-y.

[12] S. Papadakis, A. E. Kiv, H. M. Kravtsov, V. V. Osadchyi, M. V. Marienko, O. P. Pinchuk, M. P. Shyshkina, O. M. Sokolyuk, I. S. Mintii, T. A. Vakaliuk, L. E. Azarova, L. S. Kolgatina, S. M. Amelina, N. P. Volkova, V. Y. Velychko, A. M. Striuk, S. O. Semerikov, Unlocking the power of synergy: the joint force of cloud technologies and augmented reality in education, CEUR Workshop Proceedings 3364 (2023) 1–23. URL: https://ceur-ws.org/Vol-3364/paper00.pdf.

[13] K. P. Osadcha, V. V. Osadchyi, The use of cloud computing technology in professional training of future programmers, CTE Workshop Proceedings 8 (2021) 155–164. doi:10.55056/cte.229.

[14] D. V. Yatsenyak, V. P. Oleksiuk, N. R. Balyk, Study of ergonomic criteria for evaluating the software user interface, Journal of Physics: Conference Series 2288 (2022) 012005. doi:10.1088/1742-6596/2288/1/012005.

[15] S. Schez-Sobrino, D. Vallejo, C. Glez-Morcillo, M. A. Redondo, J. J. Castro-Sanchez, RoboTIC: A serious game based on augmented reality for learning programming, Multimedia Tools and Applications 79 (2020) 34079–34099. doi:10.1007/s11042-020-09202-z.

[16] N. Balyk, Y. Vasylenko, G. Shmyger, V. Oleksiuk, A. Skaskiv, Design of Approaches to the Development of Teacher's Digital Competencies in the Process of Their Lifelong Learning, CEUR Workshop Proceedings 2393 (2019) 204–219. URL: https://ceur-ws.org/Vol-2393/paper_237.pdf.

[17] M. Hidalgo, H. Astudillo, L. M. Castro, Challenges to Use Role Playing in Software Engineering Education: A Rapid Review, in: H. Florez, M. Leon (Eds.), Applied Informatics, Springer Nature Switzerland, Cham, 2024, pp. 245–260. doi:10.1007/978-3-031-46813-1_17.

[18] M. Frania, Let's Open the Locker of Creativity - How the Traditional Educational Escape Room Changed into a Virtual Puzzle Game During the COVID-19 Pandemic, in: Ł. Tomczyk (Ed.), New Media Pedagogy: Research Trends, Methodological Challenges and Successful Implementations, Springer Nature Switzerland, Cham, 2023, pp. 102–114. doi:10.1007/978-3-031-44581-1_8.

[19] M. I. M. Nadziman, H. F. M. Hanum, N. A. K. Adnan, N. M. Diah, Z. A. Bakar, Game-Based Mobile Application for Tarannum Learning, in: H. Badioze Zaman, P. Robinson, A. F.

Smeaton, R. L. De Oliveira, B. N. Jørgensen, T. K. Shih, R. Abdul Kadir, U. H. Mohamad, M. N. Ahmad (Eds.), Advances in Visual Informatics, Springer Nature Singapore, Singapore, 2024, pp. 223–233. doi:10.1007/978-981-99-7339-2_20.

[20] V. H. Andaluz, J. S. Sánchez, C. R. Sánchez, W. X. Quevedo, J. Varela, J. L. Morales, G. Cuzco, Multi-user Industrial Training and Education Environment, in: L. T. De Paolis, P. Bourdot (Eds.), Augmented Reality, Virtual Reality, and Computer Graphics, Springer International Publishing, Cham, 2018, pp. 533–546. doi:10.1007/978-3-319-95282-6_38.

[21] F. Horn, S. Vogt, S. P. Göbel, GameTULearn: An Interactive Educational Game Authoring Tool for 3D Environments, in: M. Haahr, A. Rojas-Salazar, S. Göbel (Eds.), Serious Games, Springer Nature Switzerland, Cham, 2023, pp. 384–390. doi:10.1007/978-3-031-44751-8_32.

[22] C. M. Torres-Ferreyros, M. A. Festini-Wendorff, P. N. Shiguihara-Juárez, Developing a videogame using unreal engine based on a four stages methodology, in: 2016 IEEE ANDESCON, 2016, pp. 1–4. doi:10.1109/ANDESCON.2016.7836249.

[23] T. A. Vakaliuk, O. M. Spirin, N. M. Lobanchykova, L. A. Martseva, I. V. Novitska, V. V. Kontsedailo, Features of distance learning of cloud technologies for the organization educational process in quarantine, Journal of Physics: Conference Series 1840 (2021). doi:10.1088/1742-6596/1840/1/012051.

[24] H. Żukowski, Comparison of 3D games' efficiency with use of CRYENGINE and Unity game engines, J. Comput. Sci. Inst. 13 (2019) 345–348. URL: https://ph.pollub.pl/index.php/jcsi/article/view/1330. doi:10.35784/jcsi.1330.

[25] J. Pirker, M. Bertini, M. Lux, Open source for video games: a shortlist of game engines, SIGMultimedia Rec. 12 (2022). doi:10.1145/3548570.3548579.

[26] B. Xu, T. Jing, H. Lin, J. Du, G. Zhu, CryEngine based virtual geographic environments construction, Wuhan Daxue Xuebao (Xinxi Kexue Ban)/Geomatics and Information Science of Wuhan University 42 (2017) 28 – 34. doi:10.13203/j.whugis20140768.

Application of Daubechies wavelet analysis in problems of acoustic detection of UAVs

Oleksandr Yu. Lavrynenko[1], Denys I. Bakhtiiarov[1], Bohdan S. Chumachenko[1], Oleksii G. Holubnychyi[1], Georgiy F. Konakhovych[1] and Veniamin V. Antonov[1]

[1]*National Aviation University, 1 Lubomyr Huzar Ave., Kyiv, 03058, Ukraine*

Abstract

One of the effective directions in the detection of UAVs is acoustic surveillance, the main advantage of which is the operation in passive mode, which ensures the secrecy of the applied means, and thus the safety of the operating personnel. Noise generated by the UAVs propulsion system and propeller is a significant de-masking feature. Creation and improvement of methods of detection, direction finding and recognition of small UAVs by receiving and processing of sound signals is an urgent task. When recognizing objects, the most important and problematic task is the selection of features of the acoustic signal. The selection of features affects the process of building a recognition algorithm, as well as the performance of the entire system and the quality of recognition. The use of spectral analysis allows to allocate the main features of the UAV quite effectively, such as: engine speed, the presence of harmonics of the speed, the nature of the behavior of the envelope of harmonics. A promising method for identifying the characteristic features of acoustic radiation of UAVs is Daubechies wavelet analysis. Wavelet spectrum analysis is a powerful tool for detecting and recognizing a specific type of UAV. The method provides much more informative data than simple Fourier spectral analysis. The main idea of Daubechies wavelet analysis is to decompose the studied acoustic signal by a system of Daubechies basis functions, which have special properties, in particular good localization in the time domain, which gives a significant advantage in the analysis of non-stationary acoustic signals.

Keywords

acoustic signal, UAV detection, spectrum analysis, wavelet transform, Fourier transform, Daubechies wavelet function, wavelet coefficients

CEUR-WS.org/Vol-3662/paper03.pdf

1. Introduction

Expansion of the application spheres of small unmanned aerial vehicles (UAVs) in various fields of human activity (military applications, meteorological observations, environmental protection, etc.) provides a significant economic effect. At the same time, the use of UAVs creates a number of problems associated with inadequate behavior of some UAV owners, unauthorized monitoring

CS&SE@SW 2023: 6th Workshop for Young Scientists in Computer Science & Software Engineering, February 2, 2024, Kryvyi Rih, Ukraine

oleksandrlavrynenko@gmail.com (O. Yu. Lavrynenko); bakhtiiaroff@tks.nau.edu.ua (D. I. Bakhtiiarov); bohdan.chumachenko@npp.nau.edu.ua (B. S. Chumachenko); oleksii.holubnychyi@npp.nau.edu.ua (O. G. Holubnychyi); heorhii.konakhovych@npp.nau.edu.ua (G. F. Konakhovych); veniaminas@tks.nau.edu.ua (V. V. Antonov)

0000-0002-7738-161X (O. Yu. Lavrynenko); 0000-0003-3298-4641 (D. I. Bakhtiiarov); 0000-0002-0354-2206 (B. S. Chumachenko); 0000-0001-5101-3862 (O. G. Holubnychyi); 0000-0002-6636-542X (G. F. Konakhovych); 0000-0003-2244-262X (V. V. Antonov)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

of objects and territories, etc. Accordingly, the task of UAV detection becomes relevant, which can be solved by means of active and passive radar, thermal location, video surveillance or acoustic observation systems [1, 2].

As follows from the results of studies, the total acoustic emission spectrum of a small UAV is due to harmonic random components. In known algorithms for UAV detection and direction finding, the problem is solved for a signal in a sufficiently narrow frequency band. However, the narrow-band processing of acoustic UAV signals does not allow to fully utilize the energy and information of the received signal. This becomes possible only with appropriate signal processing using wavelet analysis based on the Daubechies basis [3].

The application of spatial and temporal wavelet processing for acoustic signals of UAVs in the tasks of aircraft detection provides the expansion of the dynamic range of devices for receiving and processing signals and increasing noise immunity, which occurs due to adaptive suppression of interference in the bandwidth of the receiving device with minimal distortion of the useful signal [4]. The maximum number of suppressed interference increases, in-phase summation of acoustic signals in communication channels in the entire frequency band is provided, which allows to more fully utilize the energy of the acoustic signal of the UAV coming to the input and, consequently, allows to increase the signal-to-noise ratio at the output [5].

Thus, the implementation of wavelet algorithms for acoustic signal processing based on the Daubechies wavelet function opens up a wide range of possibilities to further improve UAV detection.

2. Literature review and problem statement

Fourier analysis is based on the statement that any 2π-periodic function can be decomposed into components, i.e., can be obtained by superposition of integer stretches of the basis function e^{ix} [6].

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx},$$

where c_n is Fourier coefficients

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} dx.$$

Fourier transform

$$\widehat{f}\ (\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} f(t) dt,$$

gives spectral information $\widehat{f}\ (\omega)$ about the acoustic signal $f(t)$ and describes its behavior in the frequency domain ω, which is very important in acoustic UAV detection [7].

When moving to the Fourier frequency domain $\widehat{f}\ (\omega)$, time information is completely lost t, which makes the Fourier spectral analysis method unsuitable for processing non-stationary acoustic signals $f(t)$, in which the determining value is the moment in time t, at which the characteristic distortions in the acoustic signal emitted by the UAV occurred [8].

In contrast to the short-time Fourier transform

$$\widehat{f}\ (\omega, t) = \int_{-\infty}^{\infty} e^{-i\omega t}\left(f(t) \cdot W\ (t)\right) dt,$$

which provides a uniform grid (figure 1) in the frequency-time domain $\widehat{f}\ (\omega, t)$ through the use of the window function $W\ (t)$, the wavelet transform has non-uniform resolution, which allows the acoustic signal of the UAV $f(t)$ to be investigated both locally and completely [9].

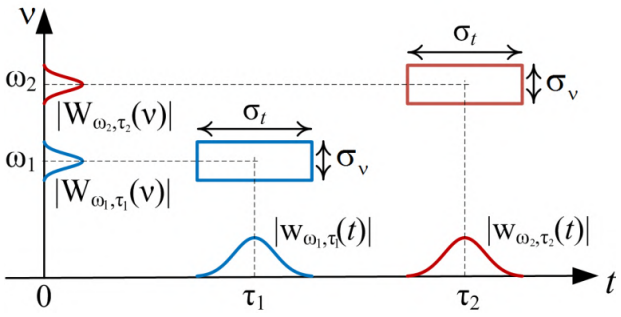


Figure 1: Time-frequency resolution of the Fourier transform.

Since the frequency ω is inversely proportional to the period T, i.e. $\omega = 1/T$, a narrower window $W\ (t)$ is required to localize the high-frequency component $\omega \to \max$ of the acoustic signal $f(t)$ and a wider window $W\ (t)$ for the low-frequency component $\omega \to \min$. The short-time Fourier transform $\widehat{f}\ (\omega, t)$ is acceptable for a signal with a relatively narrow bandwidth $\Delta\omega \to \min$, but acoustic signals $f(t)$ are not. For an acoustic signal it would be desirable to have a window $W\ (t)$, capable of changing its width with changing frequency ω [10].

Let us introduce a function $\phi \in L^2(R)$, satisfying the condition

$$\int_{-\infty}^{\infty} \frac{\left|\hat{\phi}(\omega)\right|^2}{|\omega|} d\omega < \infty,$$

and we'll call it the "base wavelet".

With respect to each basis wavelet, the wavelet transform is defined as

$$\left(\Psi_\phi f\right)(\tau, s) = |s|^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(t)\overline{\phi\left(\frac{t-\tau}{s}\right)} dt,$$

where s and τ are the scaling and shifting parameters $s, \tau \in R$; $a \neq 0$.

Then we denote

$$\phi_{\tau;s}(t) = |s|^{-\frac{1}{2}}\, \phi\left(\frac{t-\tau}{s}\right),$$

and the transformation will take the form

$$\left(\Psi_\phi f\right)(\tau, s) = \left\langle f, \phi_{\tau;s}\right\rangle.$$

If the center and radius of the window function ϕ, respectively, are equal to t^* and Δ_ϕ, then $\phi_{\tau;s}(t)$ is a window function with center $\tau + st^*$ and radius $s\Delta_\phi$. Hence, the wavelet transform localizes the signal in the time window (figure 2) [11]

$$[\tau + st^* - s\Delta_\phi, \tau + st^* + s\Delta_\phi]\,.$$

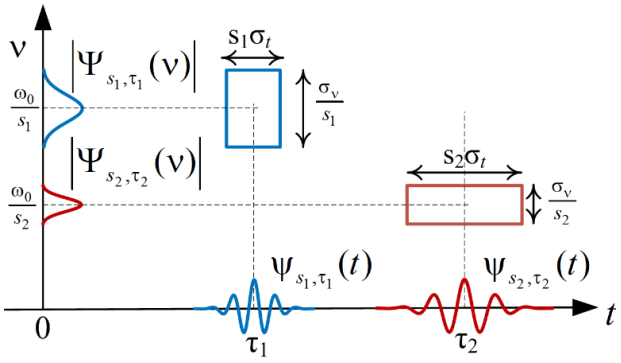


Figure 2: Time-frequency resolution of the Daubechies wavelet transform.

Thus it was shown not only that the Fourier transform is uninformative in the problems of analyzing non-stationary signals, which are acoustic signals, but also the fact that the basis of wavelet decomposition plays a major role in the effectiveness of using wavelet analysis in the task of acoustic detection of UAVs, so in this research work it is proposed to use the decomposition in Daubechies series, since this function is an orthogonal wavelet with a compact carrier computed iteratively.

3. Daubechies wavelet analysis of acoustic signals

To calculate the coefficients of the generating Daubechies wavelet filter n-th order, we need to specify only the number of zero moments of the wavelet function N, i.e., the order of the function is determined by the number of zero moments, hence $N = n$ [12].

Then the calculation of the generating Daubechies wavelet filter implies finding the coefficients of the polynomial

$$P_k = \frac{\prod_{i=-N+1}^{N}\left(\frac{1}{2} - i\right)}{\prod_{i=-N+1}^{N}(k-i)}, k = 1,\ \ldots,\ N,$$

which for all values of $k \neq i$ form a vector

$$P = \left(\ P_N\ \ 0\ \ P_{N-1}\ \ 0\ \ \ldots\ \ 0\ \ P_1\ \ 1\ \ P_1\ \ 0\ \ P_2\ \ 0\ \ \ldots\ \ 0\ \ P_N\ \right), \qquad (1)$$

length $4N - 1$.

In case $N = 1$, then all values of coefficients of polynomial $P_{1,\ \ldots,\ 4N-1}$ satisfying the condition $P_{1,\ \ldots,\ 4N-1} < 1$ define the vector

$$P = \left(\ P_1\ \ \ldots\ \ P_{2N}\ \right), \qquad (2)$$

length $2N$, whose values correspond to the coefficients of the Daubechies wavelet filter of the 1st order.

If $N > 1$ is required to compute the roots of the coefficients of the polynomial P given by the vector (1).

Then the vector of coefficients of the polynomial P is transformed into the following form

$$P = \begin{pmatrix} \frac{P_2}{P_1} & \frac{P_3}{P_1} & \dots & \frac{P_{4N-1}}{P_1} \end{pmatrix},$$

length $L = 4N - 2$.

Let's form a square matrix A of order L

$$A_L = \begin{pmatrix} -P_1 & -P_2 & \dots & -P_{L-1} & -P_L \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix},$$

where the first row of the matrix A_L defines the coefficients of the characteristic equation, which has the form

$$\lambda^L - P_1\lambda^{L-1} - P_2\lambda^{L-2} - \dots - P_{L-1}\lambda - P_L = 0, \tag{3}$$

where the roots $\lambda_{1,\,\dots,\,L}$ of this equation are the eigenvalues of the matrix A_L. The order of the square matrix A_L is always a multiple of two since $L = 4N - 2$.

Solving the equation (3) by one of the numerical methods (by the method of half division, combined, iterations, etc.), we find the roots $\lambda_{1,\,\dots,\,L}$ of this equation and thus the vector λ of eigenvalues of the matrix A_L is formed.

$$\lambda = \begin{pmatrix} \lambda_1 & \dots & \lambda_L \end{pmatrix}.$$

The values of the vector λ should be arranged in ascending order

$$\lambda = \begin{pmatrix} \lambda_{\min} & \dots & \lambda_{\max} \end{pmatrix},$$

observing the condition $|\lambda_{1,\,\dots,\,L} + 1|$, and select only those values that match the condition of the expression

$$\lambda = \begin{pmatrix} \lambda_{K+2} & \dots & \lambda_{2K} \end{pmatrix},$$

where $K = 2N - 1$, then the length of the vector λ is equal to $M = 2K - (K + 2) + 1$ values.

Let's rearrange the values of the vector λ in ascending order

$$\lambda = \begin{pmatrix} \lambda_{\min} & \dots & \lambda_{\max} \end{pmatrix},$$

complying with the condition $|\lambda_{1,\,\dots,\,M}|$.

Thus we obtain a vector λ of length M, which includes the values of the roots of $\lambda_{1,\,\dots,\,M}$ arranged in ascending order

$$\lambda = \begin{pmatrix} \lambda_1 & \dots & \lambda_M \end{pmatrix}.$$

Then all values of the roots of $\lambda_{1,\,...,\,M}$ satisfying the condition $|\lambda_{1,\,...,\,M}| < 1$ define the vector

$$\lambda = \begin{pmatrix} \lambda_1 & \dots & \lambda_H \end{pmatrix},$$

where H depends on the condition $|\lambda_{1,\,...,\,M}| > 1$, i.e., how many values of $\lambda_{1,\,...,\,M}$ are modulo greater than one.

Let's set the vector

$$O = \begin{pmatrix} O_1 & \dots & O_N \end{pmatrix},$$

where $O_{1,\,...,\,N} = -1$, since the values of the roots of $\lambda_{1,\,...,\,H}$ are complex numbers, the values of $O_{1,\,...,\,N}$ are converted to complex form, hence $O_{1,\,...,\,N} = -1.0000 + 0.0000i$.

As a result, we obtain the vector

$$\lambda = \begin{pmatrix} \lambda_1 & \dots & \lambda_H & O_1 & \dots & O_N \end{pmatrix}, \tag{4}$$

defined by the root values $\lambda_{1,\,...,\,H}$ and unit vectors $O_{1,\,...,\,N}$, of length $J = H + N$.

Then let us represent the vector λ in the form

$$\lambda = \begin{pmatrix} \lambda_1 & \dots & \lambda_J \end{pmatrix}, \tag{5}$$

equating the values of vector (4) to the notations (5).

So, having a pre-formed vector of values of roots of $\lambda_{1,\,...,\,\mathrm{J}}$ polynomial, let us calculate the vector of values of coefficients of this polynomial according to the expression

$$P_k = P_k - \lambda_j P_i, \tag{6}$$

where in cases where $j = 1,\ \dots,\ J$, then $k = 2,\ \dots,\ j+1$, $i = 1,\ \dots,\ j$, and the initial values of the coefficients correspond to the vector

$$P = \begin{pmatrix} P_1 & P_2 & \dots & P_{J+1} \end{pmatrix},$$

length $J + 1 = 2N$, where $P_1 = 1$, $P_{2,\,...,\,J+1} = 0$, since the values of the roots $\lambda_{1,\,...,\,J}$ of the polynomial are complex numbers, the values of the coefficients $P_{1,\,...,\,J+1}$ are converted to complex form, and thus $P_1 = 1.0000 + 0.0000i$, $P_{2,\,...,\,J+1} = 0.0000 + 0.0000i$.

Let us explain the recursive algorithm of expression (6) in more detail [13].

So if $j = 1,\ \dots,\ J$, where $J = 5$
we have

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

at $j = 1$

$$k = 2,\ \dots,\ j+1 = 2,\ \dots,\ 2$$

$$i = 1,\ \dots,\ j = 1,\ \dots,\ 1$$

then

$$P = \begin{pmatrix} 1 & P_2 - \lambda_1 P_1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

at $j = 2$

$$k = 2,\ \dots,\ j+1 = 2,\ \dots,\ 3$$

$$i = 1,\ \ldots,\ j = 1,\ \ldots,\ 2$$

then

$$P = \begin{pmatrix} 1 & P_2 - \lambda_2 P_1 & P_3 - \lambda_2 P_2 & 0 & 0 & 0 \end{pmatrix}$$

at $j = 3$

$$k = 2,\ \ldots,\ j+1 = 2,\ \ldots,\ 4$$

$$i = 1,\ \ldots,\ j = 1,\ \ldots,\ 3$$

then

$$P = \begin{pmatrix} 1 & P_2 - \lambda_3 P_1 & P_3 - \lambda_3 P_2 & P_4 - \lambda_3 P_3 & 0 & 0 \end{pmatrix}$$

at $j = 4$

$$k = 2,\ \ldots,\ j+1 = 2,\ \ldots,\ 5$$

$$i = 1,\ \ldots,\ j = 1,\ \ldots,\ 4$$

then

$$P = \begin{pmatrix} 1 & P_2 - \lambda_4 P_1 & P_3 - \lambda_4 P_2 & P_4 - \lambda_4 P_3 & P_5 - \lambda_4 P_4 & 0 \end{pmatrix}$$

at $j = 5$

$$k = 2,\ \ldots,\ j+1 = 2,\ \ldots,\ 6$$

$$i = 1,\ \ldots,\ j = 1,\ \ldots,\ 5$$

then

$$P = \begin{pmatrix} 1 & P_2 - \lambda_5 P_1 & P_3 - \lambda_5 P_2 & P_4 - \lambda_5 P_3 & P_5 - \lambda_5 P_4 & P_6 - \lambda_5 P_5 \end{pmatrix}.$$

Thus, according to expression (6) we obtain a vector of complex values of polynomial coefficients from which it is required to leave only the real part, and to discard the imaginary part, which will form the vector

$$P = \begin{pmatrix} P_1 & \ldots & P_{2N} \end{pmatrix}, \tag{7}$$

length $2N$, whose values correspond to the coefficients of the Daubechies wavelet filter n-th order.

Normalization of coefficients of the Daubechies wavelet filter n-th order is carried out as follows

$$P_k = S_P \frac{P_k}{\sum_{\mathrm{k=1}}^{\mathrm{2N}} \mathrm{P_k}}, \tag{8}$$

where $k = 1,\ \ldots,\ 2N$, forming the resulting vector of normalized coefficients

$$P = \begin{pmatrix} P_1 & \ldots & P_{2N} \end{pmatrix}, \tag{9}$$

such that the sum of the coefficients of $\sum_{k=1}^{2N} P_k$ will equal S_P, i.e., if $S_P = 1$, then $\sum_{k=1}^{2N} P_k = 1$ (figure 3).

Thus, at the output of the above transformations, at $N = 1$ we obtain the vector of values of coefficients of the Daubechies wavelet filter of the 1st order according to (2), and at $N > 1$ we obtain the vector of values of coefficients of the Daubechies wavelet filter of the n-order

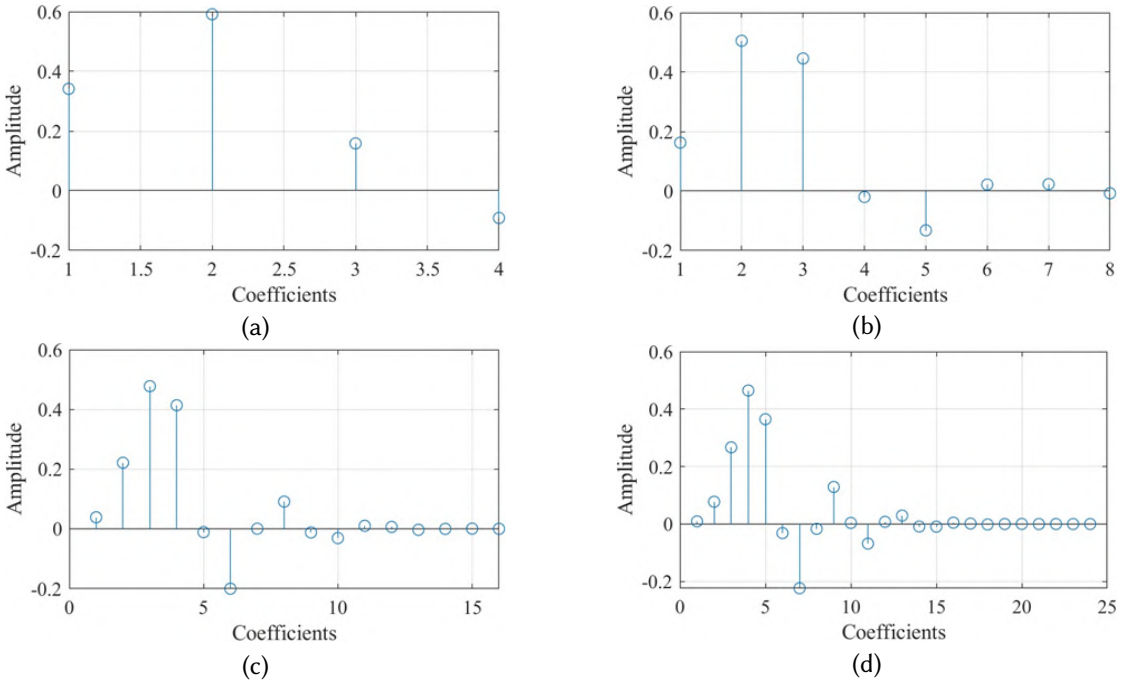


Figure 3: Coefficients of generating Daubechies wavelet filters of the 2nd (a), 4th (b), 8th (c) and 12th (d) orders.

according to (7), where in both cases the procedure of normalization of coefficients according to (8) is applied, which as a result forms the vector (9) [14].

The coefficients of the generating Daubechies wavelet filters of the 2nd, 4th, 8th and 12th orders found by the above algorithm are shown below (figure 3).

Let us calculate the coefficients of the orthogonal wavelet filters on the basis of the values of the coefficients of the generating Daubechies wavelet filter n-th order found earlier according to (9).

Thus, the coefficients of the orthogonal low-pass wavelet filter for the inverse discrete wavelet transform are defined as follows

$$R = \sqrt{2} \begin{pmatrix} P_1 & \dots & P_{2N} \end{pmatrix},$$

forming a vector

$$R = \begin{pmatrix} R_1 & \dots & R_{2N} \end{pmatrix}, \tag{10}$$

length $2N$, then the coefficients of the orthogonal low-pass wavelet filter for the direct discrete wavelet transform are defined by

$$D = \begin{pmatrix} R_{2N} & \dots & R_1 \end{pmatrix},$$

which corresponds to the inversion (10), forming the vector

$$D = \begin{pmatrix} D_1 & \dots & D_{2N} \end{pmatrix}. \tag{11}$$

The coefficients of the orthogonal high-pass wavelet filter for the inverse discrete wavelet transform are determined by computing the quadrature-mirror filter as follows

$$W = \begin{pmatrix} R_{2N} & -R_{2N-1} & R_{2N-2} & -R_{2N-3} & \dots & -R_4 & R_3 & -R_2 & R_1 \end{pmatrix},$$

forming a vector

$$W = \begin{pmatrix} W_1 & \dots & W_{2N} \end{pmatrix}, \tag{12}$$

then the coefficients of the orthogonal high-pass wavelet filter for the direct discrete wavelet transform are determined by

$$V = \begin{pmatrix} W_{2N} & \dots & W_1 \end{pmatrix},$$

which corresponds to the inversion (12), forming the vector

$$V = \begin{pmatrix} V_1 & \dots & V_{2N} \end{pmatrix}. \tag{13}$$

Thus we obtained vectors of values D and V, as well as R and W, which correspond to the coefficients of orthogonal wavelet filters of low and high frequencies for forward and inverse discrete wavelet transform, respectively [15].

As an example, let's show the coefficients of orthogonal wavelet filters based on the 8th-order Daubechies generating wavelet filter found by the above method (figure 4).

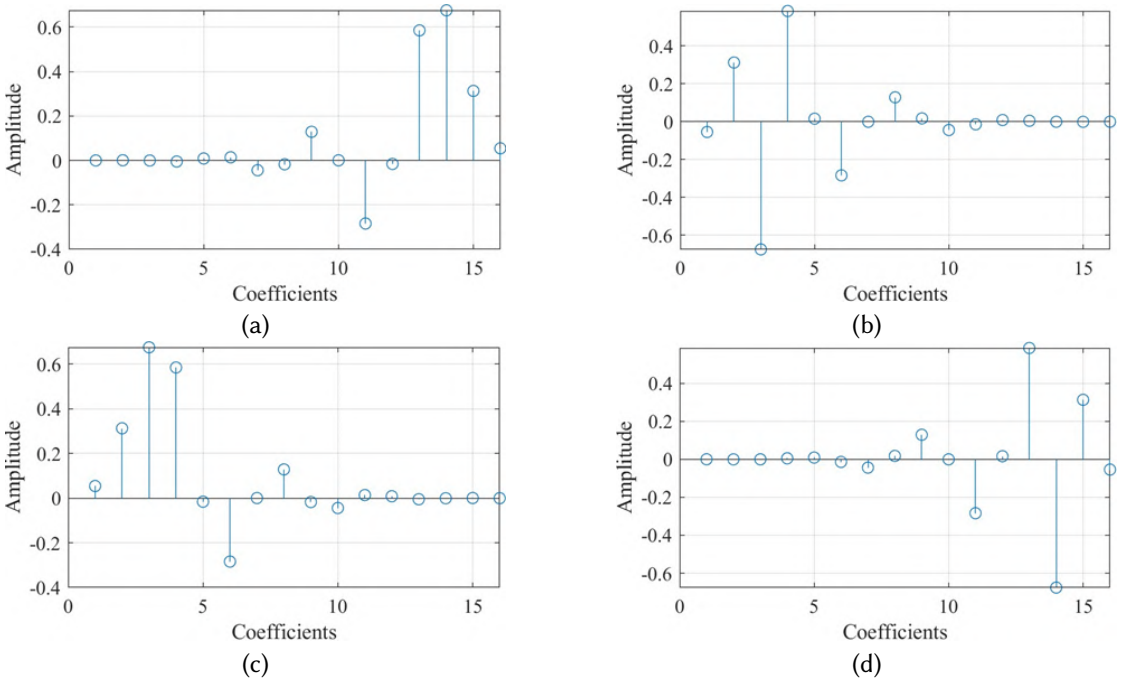


Figure 4: Coefficients of orthogonal Daubechies wavelet filters of the 8th order. (a) is low-pass decomposition filter D, (b) is high-pass decomposition filter V, (c) is low-pass reconstruction filter R, (d) is high-pass reconstruction filter W.

Then the direct discrete wavelet transform is nothing but a mathematical convolution of the values of the studied vector

$$X = \begin{pmatrix} X_1 & \dots & X_L \end{pmatrix}, \tag{14}$$

length L, with previously found vectors of coefficient values of orthogonal wavelet filters of low and high frequencies D (11) and V (13), respectively, followed by twofold thinning of $\downarrow 2$ coefficients obtained after convolution operation, thus obtaining a vector of coefficient values Z containing the low-frequency component and a vector of values Y corresponding to the high-frequency component of the studied vector X, where the formed vectors of coefficient values Z, Y are the result of this transformation [16].

Thus, the operation of mathematical convolution of the values of the investigated vector X (14) with the values of coefficients of the orthogonal low-pass wavelet filter D (11) is defined by the following expression

$$Z_k = \sum_{j=\max(1,\, k+1-2N)}^{\min(k,\, L)} \mathrm{X_j} D_i, \tag{15}$$

where $k = 1,\ \dots,\ L + 2N - 1, i = k + 1 - j,$
forming a vector of values

$$Z = \begin{pmatrix} Z_1 & \dots & Z_K \end{pmatrix},$$

where $K = L + 2N - 1$.

Let us explain expression (15) in more detail.

So, we have the vector under study

$$X = \begin{pmatrix} X_1 & \dots & X_L \end{pmatrix},$$

length $L = 8$, as well as the vector of values of coefficients of the orthogonal wavelet filter of low frequencies

$$D = \begin{pmatrix} D_1 & \dots & D_{2N} \end{pmatrix},$$

length $2N = 4$,
from where

$$k = 1,\ \dots,\ L + 2N - 1 = 1,\ \dots,\ 11$$

then according to (15)
at $k = 1, j = 1,\ \dots,\ 1, i = 1,\ \dots,\ 1$

$$Z_1 = X_1 D_1$$

at $k = 2, j = 1,\ \dots,\ 2, i = 2,\ \dots,\ 1$

$$Z_2 = X_1 D_2 + X_2 D_1$$

at $k = 3, j = 1,\ \dots,\ 3, i = 3,\ \dots,\ 1$

$$Z_3 = X_1 D_3 + X_2 D_2 + X_3 D_1$$

at $k = 4, j = 1, \ \ldots, \ 4, i = 4, \ \ldots, \ 1$

$$Z_4 = X_1 D_4 + X_2 D_3 + X_3 D_2 + X_4 D_1$$

at $k = 5, j = 2, \ \ldots, \ 5, i = 4, \ \ldots, \ 1$

$$Z_5 = X_2 D_4 + X_3 D_3 + X_4 D_2 + X_5 D_1$$

at $k = 6, j = 3, \ \ldots, \ 6, i = 4, \ \ldots, \ 1$

$$Z_6 = X_3 D_4 + X_4 D_3 + X_5 D_2 + X_6 D_1$$

at $k = 7, j = 4, \ \ldots, \ 7, i = 4, \ \ldots, \ 1$

$$Z_7 = X_4 D_4 + X_5 D_3 + X_6 D_2 + X_7 D_1$$

at $k = 8, j = 5, \ \ldots, \ 8, i = 4, \ \ldots, \ 1$

$$Z_8 = X_5 D_4 + X_6 D_3 + X_7 D_2 + X_8 D_1$$

at $k = 9, j = 6, \ \ldots, \ 8, i = 4, \ \ldots, \ 2$

$$Z_9 = X_6 D_4 + X_7 D_3 + X_8 D_2$$

at $k = 10, j = 7, \ \ldots, \ 8, i = 4, \ \ldots, \ 3$

$$Z_{10} = X_7 D_4 + X_8 D_3$$

at $k = 11, j = 8, \ \ldots, \ 8, i = 4, \ \ldots, \ 4$

$$Z_{11} = X_8 D_4.$$

Thus the vector of values is formed

$$Z = \left(\begin{array}{cccc} Z_1 & \ldots & Z_K \end{array} \right),$$

where $K = L + 2N - 1 = 11$.

Then having calculated the values of convolution coefficients $Z_{1, \ldots, K}$ according to (15), it is necessary to perform the operation of double thinning $\downarrow 2$, according to expressions

$$Z = \left(\begin{array}{ccccc} Z_2 & Z_4 & Z_6 & \ldots & Z_K \end{array} \right),$$

when K is a multiple of two, and when K is not a multiple of two

$$Z = \left(\begin{array}{ccccc} Z_2 & Z_4 & Z_6 & \ldots & Z_{K-1} \end{array} \right),$$

which in turn forms the vector

$$Z = \left(\begin{array}{ccc} Z_1 & \ldots & Z_Q \end{array} \right), \tag{16}$$

length $Q = \frac{K}{2}$ or $Q = \frac{K-1}{2}$ depending on the multiple of two K.

Thus, the found vector of coefficient values Z (16) defines the low-frequency component of the direct discrete wavelet transform of the investigated vector X.

Then to find the high-frequency component Y of the direct discrete wavelet transform of the investigated vector X, it is required to repeat the given mathematical operations (2.30 - 2.34), but respectively, for the values of the coefficients of the orthogonal high-frequency wavelet filter V (13) [17].

Thus, the operation of mathematical convolution of the values of the investigated vector X (14) with the values of coefficients of the orthogonal wavelet filter of high frequencies V (13) is defined by the following expression

$$Y_k = \sum_{j=\max(1,\, k+1-2N)}^{\min(k,\, L)} \mathrm{X_j} V_i, \tag{17}$$

where $k = 1, \ \ldots, \ L + 2N - 1, i = k + 1 - j,$
forming a vector of values

$$Y = \left(\begin{array}{cccc} Y_1 & \ldots & Y_K \end{array} \right),$$

where $K = L + 2N - 1$.

Then, having calculated the values of convolution coefficients $Y_{1,\,\ldots,\,K}$ according to (17), it is necessary to perform the operation of two-fold thinning $\downarrow 2$, according to expressions

$$Y = \left(\begin{array}{cccccc} Y_2 & Y_4 & Y_6 & \ldots & Y_K \end{array} \right),$$

when K is a multiple of two, and when K is not a multiple of two.

$$Y = \left(\begin{array}{cccccc} Y_2 & Y_4 & Y_6 & \ldots & Y_{K-1} \end{array} \right),$$

which in turn forms the vector

$$Y = \left(\begin{array}{cccc} Y_1 & \ldots & Y_Q \end{array} \right), \tag{18}$$

length $Q = \frac{K}{2}$ or $Q = \frac{K-1}{2}$ depending on the multiple of two K [18].

Then the vectors of coefficient values Z (16) and Y (18) are the result of one level of direct discrete wavelet transform, which can be written in the following form

$$\Omega = \left(\begin{array}{cccccc} Z_1 & \ldots & Z_Q & Y_1 & \ldots & Y_Q \end{array} \right),$$

then

$$\Omega = \left(\begin{array}{cccc} \Omega_1 & \ldots & \Omega_{2Q} \end{array} \right),$$

length $2Q$.

To reconstruct the studied vector X (14) by the values of wavelet coefficients $Z_{1,\,\ldots,\,Q}$ (16) and $Y_{1,\,\ldots,\,Q}$ (18), it is required to perform the operation of doubling $\uparrow 2$ coefficients, according to the expressions

$$Z = \left(\begin{array}{cccccccccc} Z_1 & 0 & Z_2 & 0 & Z_3 & 0 & \ldots & 0 & Z_{2Q-1} \end{array} \right),$$

$$Y = \left(\begin{array}{cccccccccc} Y_1 & 0 & Y_2 & 0 & Y_3 & 0 & \dots & 0 & Y_{2Q-1} \end{array} \right),$$

forming vectors

$$Z = \left(\begin{array}{ccc} Z_1 & \dots & Z_{2Q-1} \end{array} \right), \tag{19}$$

$$Y = \left(\begin{array}{ccc} Y_1 & \dots & Y_{2Q-1} \end{array} \right), \tag{20}$$

length $2Q-1$.

Then the inverse discrete wavelet transform is defined according to the expression

$$X_k = \sum_{j=\max(1,\, k+1-2N)}^{\min(k,\, 2Q-1)} \mathrm{Z_j} R_i + \sum_{j=\max(1,\, k+1-2N)}^{\min(k,\, 2Q-1)} \mathrm{Y_j} W_i, \tag{21}$$

where $k = 1,\ \dots,\ 2Q-1+2N-1$, $i = k+1-j$,
forming a vector of values

$$X = \left(\begin{array}{ccc} X_1 & \dots & X_K \end{array} \right),$$

where $K = 2Q-1+2N-1$.

Expression (21) can be characterized as the sum of two mathematical convolution of the wavelet coefficient values of $Z_{1,\,\dots,\,2Q-1}$ (19) and $Y_{1,\,\dots,\,2Q-1}$ (20) with the coefficients of the orthogonal lowpass and highpass wavelet filters $R_{1,\,\dots,\,2N}$ (10) and $W_{1,\,\dots,\,2N}$ (12), respectively [19].

From where we determine the required values $X_{1,\,\dots,\,L}$ according to the expression

$$X = \left(\begin{array}{ccc} X_{2N-1} & \dots & X_{2N-2+L} \end{array} \right),$$

then we obtain the vector

$$X = \left(\begin{array}{ccc} X_1 & \dots & X_L \end{array} \right), \tag{22}$$

length L, which is the result of the inverse discrete wavelet transform, i.e., the values of the vector X (22) are the result of the process of reconstructing the values of the studied vector X (14) by the values of the wavelet coefficients $Z_{1,\,\dots,\,Q}$ (16) and $Y_{1,\,\dots,\,Q}$ (18) [20].

4. Simulation results

The disadvantages of the Fourier transform are demonstrated in figure 5a, 5b and figure 6a, 6b.

In figure 5a and figure 6a show two harmonic components $S_1(t) = A_1 \cdot \sin(\omega_1 t)$ and $S_2(t) = A_2 \cdot \sin(\omega_2 t)$, with angular frequencies $\omega_1 = 63$ rad/s and $\omega_2 = 252$ rad/s.

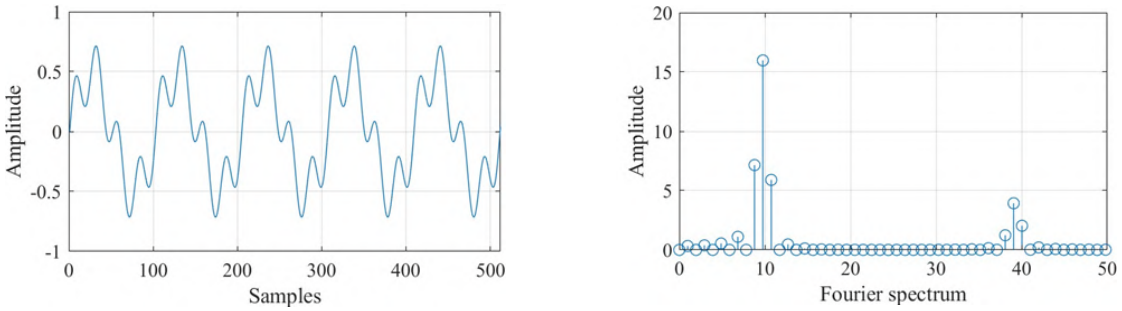
The angular frequency ω in rad/s is expressed through the frequency f in Hz, as $\omega = 2\pi f$ and $f = \frac{\omega}{2\pi}$. Based on this, $f_1 = 10$ is Hz, and $f_2 = 40$ is Hz.

The process shown in figure 5a, is an adaptive combination of two sinusoids [21] $S_1(t)$ and $S_2(t)$

$$U_1(t) = S_1(t) + S_2(t),\ t \in (0, T],$$

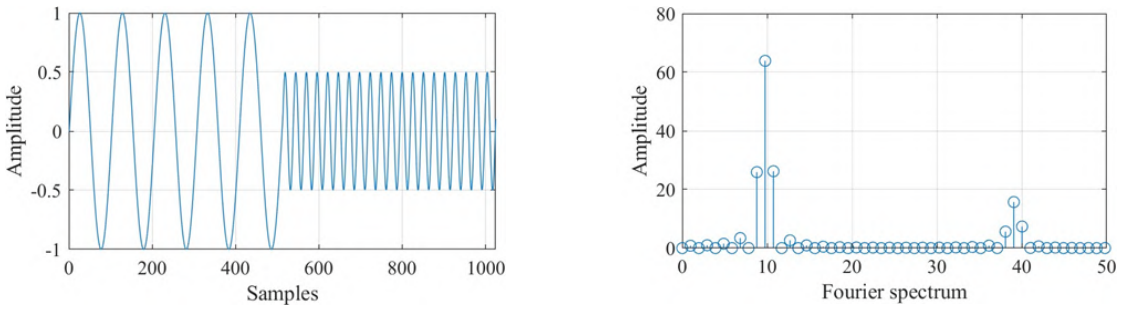
where $A_1 = 0.5$, $A_2 = 0.25$, respectively, and $T = 512$, and the process shown in figure 6a is described as follows

$$U_2(t) = \left\{ \begin{array}{l} S_1(t),\ t \in (0, t_0], \\ S_2(t),\ t \in (t_0, T], \end{array} \right.$$



(a) (b)

Figure 5: Fourier transform of the signal $U_1(t)$. (a) is signal $U_1(t)$; (b) is Fourier transform spectrum of the signal $U_1(t)$.



(a) (b)

Figure 6: Fourier transform of the signal $U_2(t)$. (a) is signal$U_2(t)$; (b) is Fourier transform spectrum of the signal $U_2(t)$.

where $A_1 = 0.5$, $A_2 = 0.25$ respectively and $t_0 = 512$ where $T = 1024$.

Outside the interval $(0, T]$, the functions $U_1(t)$ and $U_2(t)$ are 0.

As a result of the Fourier transform of the signals $U_1(t)$ and $U_2(t)$, we obtained poorly distinguishable spectral images, which are shown in figure 5b and figure 6b.

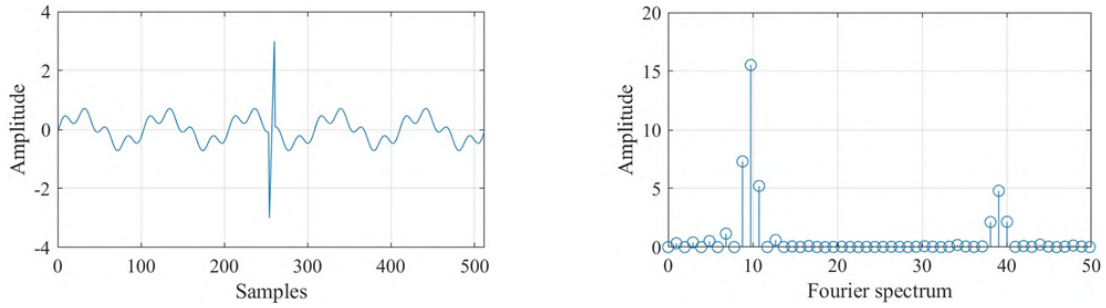
The following example also shows the low information content of the Fourier transform. The signal presented in figure 7a, the signal $U_3(t)$ in the vicinity of $t = 253 : 260$ contains a short pulse $I(t)$ (anomaly) [22], where $t \in (-3, 3]$

$$U_3(t) = \begin{cases} U_1(t), & t \in (0, t_1], \\ I(t), & t \in (t_1, t_2], \\ U_1(t), & t \in (t_2, T], \end{cases}$$

where $t_1 = 253$, $t_2 = 260$.

The Fourier transform made it possible to clearly distinguish two harmonic components of the signal, and the spectral components of the anomaly, as expected, were distributed along the entire frequency axis.

In figure 5, figure 6, figure 7 showed specific examples of the disadvantages of the Fourier transform that can be overcome by using the wavelet transform.

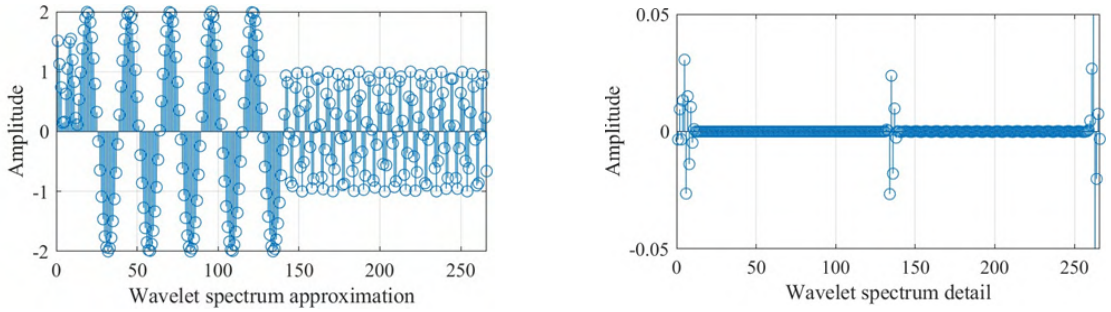


(a) (b)

Figure 7: Fourier transform of the signal $U_3(t)$. (a) is signal $U_3(t)$; (b) is Fourier transform spectrum of the signal $U_3(t)$.

It should be noted that the above Fourier transform spectra contain all the information about the input signals. This information is distributed in the phase and amplitude values of all spectral components. The input acoustic signals can be fully recovered after the inverse Fourier transform.

The advantages of the wavelet transform are demonstrated in figure 8a, 8b, and figure 9a, 9b.

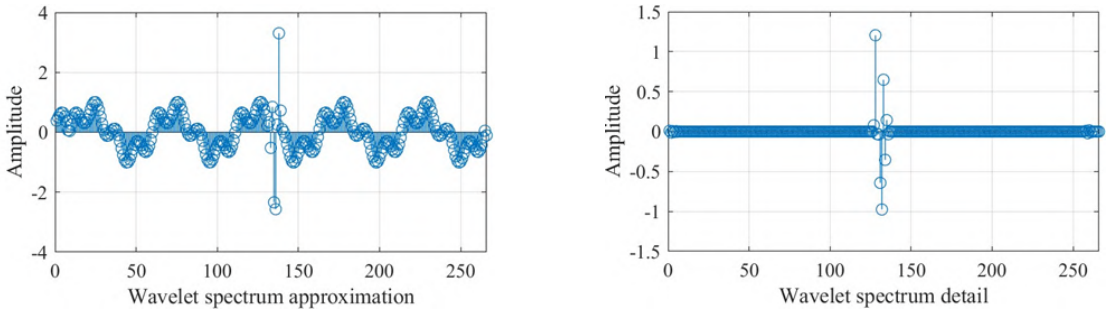


(a) (b)

Figure 8: Wavelet transform of the signal $U_2(t)$. (a) is approximation coefficients of the wavelet transform of the signal $U_2(t)$; (b) is detail coefficients of the wavelet transform of the signal $U_2(t)$.

5. Discussion

As a spectral analysis of a noisy acoustic signal, it was proposed to use a wavelet transform based on the Daubechies wavelet function. This transformation has advantages over the Fourier transform, as it is adaptive to obtain a set of informative acoustic features for UAV recognition, which will keep the classification at a sufficiently high level. As a result of the first step of the wavelet transform, the time resolution is halved, since only half of the samples characterize the entire acoustic signal. However, the frequency resolution is doubled, as the signal now occupies half the frequency band and the uncertainty is reduced. This procedure, known as subband coding, is repeated further and the wavelet coefficients at the output of the low-pass filter are



(a) (b)

Figure 9: Wavelet transform of the signal $U_3(t)$. (a) is approximation coefficients of the wavelet transform of the signal $U_3(t)$; (b) is detail coefficients of the wavelet transform of the signal $U_3(t)$.

fed to the same processing circuit, and the wavelet coefficients at the output of the high-pass filter are considered the resultant wavelet coefficients.

The most significant frequencies of the input acoustic signal will be displayed as large amplitudes of wavelet coefficients that characterize the corresponding frequency range. Small values of wavelet coefficients mean low energy of the corresponding frequency bands in the acoustic signal. These coefficients can be set to zero without significant signal distortion, which is very promising in the formation of acoustic signal recognition features for UAV detection.

A wavelet transform is a decomposition of an acoustic signal into a system of wavelet functions, each of which is a shifted and scaled copy of one function - the parent wavelet. Usually, the parameter that determines the choice of the type of mother wavelet is the external similarity of the signal under study and the transformation function. Based on this, it is advisable to use Daubechies wavelets as the mother wavelet function for processing acoustic signals.

This is one of the most famous wavelets and its main properties are as follows:

1) the functions have a finite number of zero values, i.e., the Daubechies wavelet system has the properties of smoothness and moment exclusion;
2) the functions have the properties of carrier compactness (rapidly increasing and rapidly decreasing) and orthogonality, which makes it possible to accurately restore the acoustic signal;
3) wavelets have both a wavelet function and a scaling function, which makes it possible to perform multiple-scale and fast wavelet analysis.

Functions on the same scale and on different scales are orthogonal. Note that the property of orthogonality allows us to obtain independent information at different scales, and normalization ensures that the value of information is preserved at different stages of the transformation. Among the disadvantages is the asymmetry of the Daubechies wavelet.

In acoustic signal processing tasks for UAV detection by noise, due to the unique sound characteristics of UAVs, the requirements imposed on the shape of wavelet function spectra are quite high, which leads to the use of a large number of zero moments (10-15 zero moments). Daubechies wavelets of length L have $= L/2$ zero moments. However, it should be remembered that the number of zero moments determines the length of wavelet functions and, therefore, the

speed of the algorithm for calculating the wavelet transform. In the classical Daubechies design, the length of the filters is $L = 2$, where M is the number of zero moments. All Daubechies wavelet functions have a compact carrier.

It is easy to see that the smoothness of wavelets increases as their order increases. At the same time, the frequency of oscillations increases. These wavelets have a characteristic asymmetry, namely the rise of the function is stretched compared to the decay.

The main problem when working with a wavelet transform is the problem of choosing the most appropriate wavelet. The choice of a particular family of wavelets is dictated by the application tasks and the type of information about the signal that needs to be maximally detected (recognized). There are no hard and fast rules, but it is best to choose a wavelet so that it belongs to the same class of functions as the signal being analyzed. If the original function can be approximated by a polynomial, then the number of zero moments of the wavelet should be approximately equal to the degree of the polynomial. The number of zero moments is more important to achieve higher information content of wavelet coefficients, which increases with a large number of zero moments.

6. Conclusion

This research paper is devoted to the wavelet analysis of acoustic signals of UAVs, which can improve the efficiency of aircraft detection algorithms. The problem of spatial and temporal wavelet processing of the received UAV acoustic signal by the criterion of maximum useful signal-to-noise ratio on the basis of Daubechies wavelet basis is considered.

The necessary mathematical relations determining the sequence of processing of the received acoustic signal on the basis of wavelet analysis using the Daubechies decomposition basis are obtained. The vector of optimal weighted Daubechies wavelet coefficients is formed in accordance with one of the known criteria of optimality of spatial and temporal processing, for example, in accordance with the criterion of maximum signal-to-noise ratio.

The obtained simulation results reflect the effectiveness of the spatio-temporal wavelet method for processing acoustic signals of UAVs in the Daubechies decomposition basis compared to the less effective Fourier basis, and as a consequence, indicate its applicability for solving problems related to the detection of UAVs by the acoustic method.

Further scientific research, continuing this topic, will be related to the construction of primary acoustic features for UAV recognition. Acoustic noise emitted by a UAV is a realization of a broadband random process, the description of which can be given by an energy wavelet spectrum. Therefore, the information attributes of acoustic recognition of UAVs can serve as estimates of spectral wavelet coefficients determined from a discrete realization containing a given number of samples. The transition to secondary information features is carried out by constructing the covariance matrix of spectral wavelet coefficients and its diagonalization. After the calculations, the set of acoustic signs of UAV recognition, which came to the input of the system, corresponds to some class, if the average value of the similarity coefficient for all pairs of vectors is greater than a certain threshold value. The conducted theoretical studies allow us to develop a module for the formation of a collection of acoustic recognition features of UAVs and a module that implements the decision-making rule for the classification of feature vectors.

References

[1] J. Sadovskis, A. Aboltins, Modern methods for UAV detection, classification, and tracking, in: 2022 IEEE 63th International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON), 2022, pp. 1–7. doi:10.1109/RTUCON56726.2022.9978860.

[2] A. R. Petrosian, R. V. Petrosyan, I. A. Pilkevych, M. S. Graf, Efficient model of PID controller of unmanned aerial vehicle, Journal of Edge Computing 2 (2023) 104–124. doi:10.55056/jec.593.

[3] S. Pohasii, R. Korolov, N. Dzheniuk, A. Jammine, T. Andriushchenko, T. Milevska, Decision Making in Managing the Choice of UAV Threat Detection Systems in the Protection of Critical Infrastructure Facilities, in: 2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), 2023, pp. 1–6. doi:10.1109/HORA58378.2023.10155776.

[4] V. Kartashov, V. Oleynikov, I. Koryttsev, S. Sheiko, O. Zubkov, S. Babkin, I. Selieznov, Use of Acoustic Signature for Detection, Recognition and Direction Finding of Small Unmanned Aerial Vehicles, in: 2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), 2020, pp. 1–4. doi:10.1109/TCSET49122.2020.235458.

[5] H. Deng, H. Wang, Y. Ji, X. Sun, X. Cao, Research on UAV Detection and Classification Based on Time-Frequency Analysis and Support Vector Machine, in: 2023 8th International Conference on Computer and Communication Systems (ICCCS), 2023, pp. 539–543. doi:10.1109/ICCCS57501.2023.10151387.

[6] K. Bremnes, R. Moen, S. R. Yeduri, R. R. Yakkati, L. R. Cenkeramaddi, Classification of UAVs Utilizing Fixed Boundary Empirical Wavelet Sub-Bands of RF Fingerprints and Deep Convolutional Neural Network, IEEE Sensors Journal 22 (2022) 21248–21256. doi:10.1109/JSEN.2022.3208518.

[7] Y. Xue, G. Jin, T. Shen, L. Tan, N. Wang, J. Gao, L. Wang, SmallTrack: Wavelet Pooling and Graph Enhanced Classification for UAV Small Object Tracking, IEEE Transactions on Geoscience and Remote Sensing 61 (2023) 1–15. doi:10.1109/TGRS.2023.3305728.

[8] S. D. Thepade, P. Bidwai, Iris recognition using fractional coefficients of transforms, Wavelet Transforms and Hybrid Wavelet Transforms, in: 2013 International Conference on Control, Computing, Communication and Materials (ICCCCM), 2013, pp. 1–5. doi:10.1109/ICCCCM.2013.6648921.

[9] S. D. Thepade, J. H. Dewan, S. S. Erandole, S. R. Jadhav, Extended performance comparison of self mutated hybrid wavelet transforms in image compression with hybrid wavelet transforms & orthogonal transforms, in: 2015 Global Conference on Communication Technologies (GCCT), 2015, pp. 317–322. doi:10.1109/GCCT.2015.7342675.

[10] I. Sornsen, C. Suppitaksakul, R. Kitpaiboontawee, Partial Discharge Signal Detection in Generators Using Wavelet Transforms, in: 2021 International Conference on Power, Energy and Innovations (ICPEI), 2021, pp. 195–198. doi:10.1109/ICPEI52436.2021.9690682.

[11] L. R. Soares, H. M. De Oliveira, R. J. S. Cintra, Signal Analysis Using Fourier-like Wavelets, in: 2006 IEEE/PES Transmission & Distribution Conference and Exposition: Latin America, 2006, pp. 1–6. doi:10.1109/TDCLA.2006.311514.

[12] X. Zhe, Z. Di, De-noising of spectral signal based on stationary wavelet transform, in: 2011 International Conference on Electric Information and Control Engineering, 2011, pp. 3098–3101. doi:10.1109/ICEICE.2011.5777848.

[13] M. Omachi, S. Omachi, Fast calculation of continuous wavelet transform using polynomial, in: 2007 International Conference on Wavelet Analysis and Pattern Recognition, volume 4, 2007, pp. 1688–1691. doi:10.1109/ICWAPR.2007.4421725.

[14] V. Nichoga, V. Pavlysh, Y. Romanyshyn, Features of use wavelet transforms for processing and analysis of rail fault detection signals, in: 2010 International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET), 2010, pp. 295–295.

[15] D. I. Bakhtiiarov, G. F. Konakhovych, O. Y. Lavrynenko, An Approach to Modernization of the Hat and COST 231 Model for Improvement of Electromagnetic Compatibility in Premises for Navigation and Motion Control Equipment, in: 2018 IEEE 5th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC), 2018, pp. 271–274. doi:10.1109/MSNMC.2018.8576260.

[16] O. Veselska, O. Lavrynenko, R. Odarchenko, M. Zaliskyi, D. Bakhtiiarov, M. Karpinski, S. Rajba, A Wavelet-Based Steganographic Method for Text Hiding in an Audio Signal, Sensors 22 (2022) 5832. doi:10.3390/s22155832.

[17] O. Lavrynenko, R. Odarchenko, G. Konakhovych, A. Taranenko, D. Bakhtiiarov, T. Dyka, Method of Semantic Coding of Speech Signals based on Empirical Wavelet Transform, in: 2021 IEEE 4th International Conference on Advanced Information and Communication Technologies (AICT), 2021, pp. 18–22. doi:10.1109/AICT52120.2021.9628985.

[18] D. Bakhtiiarov, G. Konakhovych, O. Lavrynenko, Protected system of radio control of unmanned aerial vehicle, in: 2016 4th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC), 2016, pp. 196–199. doi:10.1109/MSNMC.2016.7783141.

[19] R. Odarchenko, O. Lavrynenko, D. Bakhtiiarov, S. Dorozhynskyi, V. A. O. Zharova, Empirical Wavelet Transform in Speech Signal Compression Problems, in: 2021 IEEE 8th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T), 2021, pp. 599–602. doi:10.1109/PICST54195.2021.9772156.

[20] V. Kuzmin, M. Zaliskyi, O. Holubnychyi, O. Lavrynenko, Empirical Data Approximation Using Three-Dimensional Two-Segmented Regression, in: 2022 IEEE 3rd KhPI Week on Advanced Technology (KhPIWeek), 2022, pp. 1–6. doi:10.1109/KhPIWeek57572.2022.9916335.

[21] T. Guo, T. Zhang, E. Lim, M. López-Benítez, F. Ma, L. Yu, A Review of Wavelet Analysis and Its Applications: Challenges and Opportunities, IEEE Access 10 (2022) 58869–58903. doi:10.1109/ACCESS.2022.3179517.

[22] E. A. Martinez-Ríos, R. Bustamante-Bello, S. Navarro-Tuch, H. Perez-Meana, Applications of the Generalized Morse Wavelets: A Review, IEEE Access 11 (2023) 667–688. doi:10.1109/ACCESS.2022.3232729.

Data processing method for multimodal distribution parameters estimation

Oleksandr V. Solomentsev[1], Maksym Yu. Zaliskyi[1], Denys I. Bakhtiiarov[1] and Bohdan S. Chumachenko[1]

[1]National Aviation University, 1 Liubomyra Huzara Ave., Kyiv, 03058, Ukraine

Abstract

The increase in the amount of data makes it necessary to develop new methods of their processing. In telecommunications and radio engineering, this trend is associated with the complication of signals for the transmission of messages and an increase in the measurement parameters of both the equipment itself and the processes of its operation. During the operation of information transmission systems, the task of evaluating the parameters of the received signals, which are usually affected by interference, is important. This paper considers the problems of synthesis and analysis of a data processing method for estimating the parameters of multimodal distributions. The problem of synthesis is considered on the example of the trimodal probability density function of the sample population, which includes chaotic impulse noise of positive and negative polarity. The problem of analysis is solved on the basis of statistical simulation.

Keywords

Data processing, estimation, method of moments, method of quantiles, synthesis and analysis

1. Introduction

The development of Industry 4.0 is accompanied by an increase in data volumes in all its systems [1, 2]. The capabilities of monitoring systems and computing systems make it easy to collect, store and process these data [3, 4]. Intelligent data processing technologies give the opportunity to implement the principles of data-driven decision-making [5], which significantly increases the efficiency of using equipment for its intended purpose.

Information and measurement systems with use of statistical data processing technologies solve problems of testing hypotheses, detection, estimation and measurement of distribution parameters, filtration and extrapolation, pattern recognition, and others [6]. To ensure the efficient functioning of measured data processing structures, it is advisable to have a priori information about the parameters that characterize the distribution $f(n)$ of the noise component

CS&SE@SW 2023: 6th Workshop for Young Scientists in Computer Science & Software Engineering, February 2, 2024, Kryvyi Rih, Ukraine

avsolomentsev@ukr.net (O. V. Solomentsev); maximus2812@ukr.net (M. Yu. Zaliskyi); bakhtiiaroff@tks.nau.edu.ua (D. I. Bakhtiiarov); body21033@gmail.com (B. S. Chumachenko)
https://tks.nau.edu.ua/vikladatskij-sklad/solomentsev-oleksandr-vasylovych/ (O. V. Solomentsev); https://tks.nau.edu.ua/vikladatskij-sklad/zaliskyj-m-yu/ (M. Yu. Zaliskyi); https://tks.nau.edu.ua/vikladatskij-sklad/bahtiyarov-denys-ilshatovych/ (D. I. Bakhtiiarov); https://tks.nau.edu.ua/chumachenko-bohdan-sergijovych/ (B. S. Chumachenko)
0000-0002-3214-6384 (O. V. Solomentsev); 0000-0002-1535-4384 (M. Yu. Zaliskyi); 0000-0003-3298-4641 (D. I. Bakhtiiarov); 0000-0002-0354-2206 (B. S. Chumachenko)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

CEUR-WS.org/Vol-3662/paper16.pdf

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

[7]. If such information is missing, then it is necessary to have estimates of the parameters of the probability density function (PDF) $f(n)$ [8].

The analysis showed that it is quite difficult to obtain an analytical solution to the problem of synthesizing an algorithm for estimating PDF parameters within the framework of any of the known methods of estimation theory if the type of PDF is non-Gaussian [9, 10]. Therefore, this paper considers the problem of synthesizing a suboptimal method for estimating PDF parameters based on a combination of using the method of moments and the method of quantiles.

2. Literature review and problem statement

During the operation of radioelectronic and telecommunication systems, control actions are formed to maintain the efficiency of using the equipment for its functional purpose [11, 12]. Control actions are formed based on the results of monitoring the condition of equipment, components of the operation system, electromagnetic environment, and others [13]. As a rule, information signals, parameters and data that characterize monitoring results are stochastic [14]. In radioelectronic and telecommunication systems, data can be associated with the trends of changes in defining parameters, reliability indicators, and information signals for transmitting messages [15, 16].

While measuring the defining parameters and reliability indicators, control and measuring equipment is used, which can be located close to or remote from the equipment [17, 18]. In this case, interference is possible, which is observed especially when monitoring the electromagnetic environment [19, 20]. Data transmission channels may also be subject to interference influence [21]. Interference distorts objective data about the state of radioelectronic and telecommunication systems [22, 23]. Data processing algorithms must be developed on the principles of adaptation and readiness to process data with noise [24, 25].

For adaptation, it is necessary to estimate the interference parameters, and for this we need appropriate estimation algorithms [26].

The literature presents a wide variety of methods for estimating distribution parameters. Among these methods are [6, 7, 27]:

1. Maximum likelihood method.
2. Method of moments.
3. Method of maximum posterior probability.
4. Method of quantiles.
5. Heuristic methods and others.

Let us consider the generalized statement of the problem of this paper. The block diagram of data processing includes a number of algorithms $\overrightarrow{Algorithms}(\overrightarrow{Data}/signal, noise)$. Knowledge of signal and noise patterns is the key to high-quality and efficient data processing for decision-making. In this case, we can consider a generalized operator that generates efficiency estimates and is associated with the functioning of data processing algorithms

$$\overrightarrow{Efficiency} = \Theta(\overrightarrow{Algorithms}(\overrightarrow{Data}/signal, noise)). \quad (1)$$

Evaluation and optimization of the efficiency is a complex task, so in this paper we will consider the problem of estimating the noise parameters for a given signal-interference situation.

3. Synthesis of data processing method for estimating the parameters of multimodal distributions

While solving problems of statistical data processing, a following model is often used for describing samples of measurement information

$$y(t) = s(t) + n(t).$$

where $y(t)$ is observable process; $s(t)$ is signal component, which reflects the objective process of changing the properties of the phenomenon under study; $n(t)$ is noise component, which is caused by errors in control and measuring equipment and the presence of interference in the measuring circuits.

Analysis of measurement data shows that the noise component may include chaotic pulsed noise of both positive and negative polarity relative to the nominal level. For this case, the noise component can be characterized by the PDF of the following form

$$f(n) = (1 - p_1 - p_2)N(m_1(n) = U_+ = U_- = 0, \sigma(n/U_+ = U_- = 0)) +$$
$$+ p_1 N(m_1(n) = U_+, \sigma(n/U_+ = U_- = 0)) + p_2 N(m_1(n) = U_-, \sigma(n/U_+ = U_- = 0)), \quad (2)$$

where $N(m_1(n) = U_+ = U_- = 0, \sigma(n/U_+ = U_- = 0))$ is normal PDF of sample values in case when chaotic pulsed noise is absent; $N(m_1(n) = U_+, \sigma(n/U_+ = U_- = 0))$ is normal PDF of sample values in case when chaotic pulsed noise is occurred with positive amplitude U_+ and average value of probability of pulsed noise presence p_1; $N(m_1(n) = U_-, \sigma(n/U_+ = U_- = 0))$ is normal PDF of sample values in case when chaotic pulsed noise is occurred with negative amplitude U_- and average value of probability of pulsed noise presence p_2; $\sigma(n/U_+ = U_- = 0)$ is standard deviation for noise component for those values of the sample for which chaotic pulsed noise is absent; $m_1(n)$ is expected value of noise component.

As can be seen, equation (2) is a weighted sum of three normal distributions. In this case, the noise component is a non-stationary random process, which in general depends on time. Therefore, to simplify mathematical calculations, we will make the assumption that the mixture of the noise component is stationary. Then the data sample is homogeneous, and its mathematical expectation and standard deviation do not depend on time. It should be noted that for the example under consideration, five parameters need to be estimated, namely $\sigma(n/U_+ = U_- = 0), U_+, U_-, p_1, p_2$. In addition, we will assume that between the parameters $\sigma(n/U_+ = U_- = 0), U_+, U_-$ the following relationship exists

$$|U_+| \geqslant 3\sigma(n),$$

$$|U_-| \geqslant 3\sigma(n).$$

The procedure for synthesizing a method for estimating PDF parameters consists of two stages. The estimation algorithm is based on a fixed and known sample size m. At the same time, we believe that when forming a training sample $\overrightarrow{y_m}$, there is no signal component in the measured process. Then the PDF of the mixture will coincide with the PDF of the noise (2). The estimation algorithm splits the original sample into two parts. In this case, the first part contains samples of the positive region, and the second contains negative values of samples.

In accordance with this assumption, when comparing samples y_i with a zero threshold, two samples are formed y_{m+} and y_{m-}. Let's denote the sample y_{m+} values as y_{i1}, and the sample y_{m-} values as y_{i2}.

The sample size m_+ corresponds to the situation when $y_i > 0$; the sample size c m_- corresponds to the situation when $y_i \leqslant 0$. In general, the equation $m_+ + m_- = m$ is correct.

For the estimation method, two pairs of sampling thresholds are presented, namely V_{1+}, V_{2+} and V_{1-}, V_{2-}. A possible view of the PDF $f(n)$ and the location of the thresholds $V_{1+}, V_{2+}, V_{1-}, V_{2-}$ is shown schematically in figure 1.

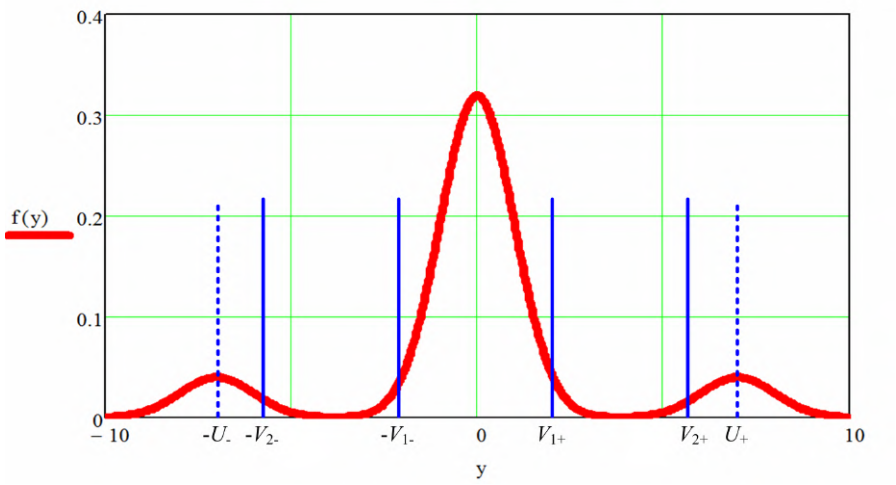


Figure 1: The view of the PDF $f(n)$ and the location of the thresholds for initial sample.

The thresholds V_{1+}, V_{2+} can be choosen using following conditions:

$$\begin{cases} V_{1+} < V_{2+}, \\ \sigma(n) < V_{1+} < 2\sigma(n)), \\ 3\sigma(n) < V_{2+} < U_+. \end{cases} \tag{3}$$

This means that the threshold V_{1+} should not exceed the samples of that part of the chaotic impulse noise, the mathematical expectation of which is zero, and the threshold V_{2+}, in addition, should not exceed the samples of that part of the chaotic impulse noise, the mathematical expectation of which is $m_1(y) = m_1(n) = U_+$. Similar considerations for choosing thresholds V_{1-}, V_{2-} take place for samples $y_i \leqslant 0$.

Sample populations $\overrightarrow{y_1}$ and $\overrightarrow{y_2}$ have a probabilistic description that is different from that used for the noise term in equation (2).

In particular case, for training sample $\overrightarrow{y_1}$ when $0 < y_i < \infty$ one-dimensional PDF of y_{i1} has

the following form

$$\begin{cases} f(y_1/p_{1+}, U_+, \sigma(y_1)) = (1-p_{1+})\frac{2}{\sqrt{2\pi}\sigma(y_1)} exp(-\frac{y_1^2}{2\sigma^2(y_1)}) + \\ \qquad +p_{1+}\frac{1}{\sqrt{2\pi}\sigma(y_1)} exp(-\frac{(y_1-U_+)^2}{2\sigma^2(y_1)}), \\ \qquad \overrightarrow{y_1} > 0, \\ \qquad p_{1+} = \frac{p_1}{A_1}, \\ \qquad A_1 = \int_0^\infty f(y/p_1, p_2, U_+, U_-, \sigma(y))dy, \end{cases} \tag{4}$$

where p_{1+} is the average probability of the appearance of chaotic pulse noise of positive polarity with an average amplitude U_+ for the samples y_{1i} that satisfy the condition $y_{1i} > 0$; A_1 is normalization factor, which takes into account the truncated nature of the PDF; $\sigma(y_1)$ is standard deviation of sample y_{1i} that coincides with standard deviation $\sigma(y)$ of initial PDF (2).

The PDF (4) is a weighted sum of two distributions: 1) truncated normal and 2) normal, which corresponds to the chaotic pulse noise of positive polarity with the average amplitude U_+. Taking into account the form of the original PDF (2), the truncated normal distribution corresponds to the situation of the absence of chaotic pulse noise of positive polarity.

The number of unknown parameters of PDF (4) is equal to three, not counting the normalization factor A_1, the estimate of which can be obtained using the quantile method in accordance with the following formulas:

$$A_1^* = \frac{1}{m}\sum_{i=1}^{m}\xi_i;$$

$$\xi_i = \begin{cases} 1, y_i > 0, \\ 0, y_i \leqslant 0; \end{cases}$$

$$m_+ = \sum_{i=1}^{m}\xi_i.$$

Taking into account the conditions for setting discretization thresholds (2), we will obtain a system of three equations for unknown parameters $p_{1+}, U_+, \sigma(y_1)$ using two estimation methods: the method of moments and the method of quantiles. In accordance with the quantile method, we equate the sample estimate $h_1^*(0 < y_{1i} \leqslant V_{1+})$ of the probability of not exceeding the threshold V_{1+} to a theoretically determined value $h_1(0 < y_{1i} \leqslant V_{1+})$, taking into account PDF (4).

The second equation of the system is obtained after equating the values of $h_2^*(0 < y_{1i} \leqslant V_{2+})$ and $h_2(0 < y_{1i} \leqslant V_{2+})$. For the third equation of the system, we use the method of moments within the framework of the first initial moment of the random variable y_{1i}. The system of equations will take the following form:

$$\begin{cases} h_{1+}^* = \int_0^{V_{1+}} (1-p_{1+})\frac{1}{\sqrt{2\pi}\sigma(y_1)} \exp\left(-\frac{y_1^2}{2\sigma^2(y_1)}\right) dy_1, \\ h_{2+}^* = (1-p_{1+}) + \int_0^{V_{2+}} \frac{p_{1+}}{\sqrt{2\pi}\sigma(y_1)} \exp\left(-\frac{(y_1-U_+)^2}{2\sigma^2(y_1)}\right) dy_1, \\ m_1^*\left(y_1/0 < y_1 < \infty\right) = (1-p_{1+})\sqrt{\frac{2}{\pi}}\sigma\left(y_1\right) + p_{1+}U_+, \end{cases} \tag{5}$$

where

$$m_1^*(y_1) = \frac{\sum_{i=1}^{m+} y_{1i}}{m_+}; h_{1+}^* = \frac{\sum_{i=1}^{m+} \xi_i^/}{m_+}; h_{2+}^* = \frac{\sum_{i=1}^{m+} \xi_i^{//}}{m_+};$$

$$\xi_i^/ = \begin{cases} 1, 0 < y_i \le V_{1+}, \\ 0, y_i > V_{1+}; \end{cases}$$

$$\xi_i^{//} = \begin{cases} 1, 0 < y_i \le V_{2+}, \\ 0, y_i > V_{2+}. \end{cases}$$

Note that the first equation in the system (5) is obtained under the condition

$$\begin{cases} \beta_2 \le \beta_1, \\ \beta_1 = (1 - p_{1+}) \int_0^{V_{1+}} \frac{1}{\sqrt{2\pi}\sigma(y_1)} \exp\left(-\frac{y_1^2}{2\sigma^2(y_1)}\right) dy_1, \\ \beta_2 = p_1 \int_0^{V_{1+}} \frac{1}{\sqrt{2\pi}\sigma(y_1)} \exp\left(-\frac{(y_{1+}-U_+)^2}{2\sigma^2(y_1)}\right) dy_1. \end{cases}$$

When calculating integrals in the system of equations (5), we can utilize the Laplace form of the probability integral. In this case, we employ a linear approximation of the Laplace form probability integral in the following way

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{z^2}{2}} dz = k(x + a) \, for - a \le x \le a. \tag{6}$$

The parameter k in equation (6) is chosen in such a way that the condition $f(x = 0) = 0.5$ is satisfied when $x = 0$. In particular, for $a = 2$, the parameter $k = 1/4$, and for $a = 3$, the parameter $k = 1/6$.

In general, linear approximations of the probability integral in Laplace form can have different values of the parameter k in equation (6). Therefore, we assume that in formula (5) when determining h_{1+}^* and h_{2+}^* we use the following approximation

$$\begin{cases} \Phi(y_1) = k_1(y_1 + a_1), \\ \Phi(y_1) = k_2(y_1 + a_2). \end{cases}$$

The solution of the system of equations (5) will yield the following expressions for determining the desired parameters of PDF (4)

$$\begin{cases} \sigma^*(y_1) = \left[-E_2 \pm \left(E_2^2 - 4E_1E_3\right)^{1/2}\right](2E_1)^{-1}, \\ p_{1+}^* = 1 - h_{1+}^* \sigma^*(y_1)(2k_1V_{1+})^{-1}; \\ U_+^* = \left[m_1^*(y_1) - (1 - p_1^*)\sigma^*(y_1)\sqrt{\frac{2}{\pi}}\right](p_1^*)^{-1}, \end{cases} ; \tag{7}$$

where

$$E_1 = h_1^* \left(1 + k_2\sqrt{\frac{2}{\pi}} - k_2a_2\right);$$

$$E_2 = 2k_2a_2k_1V_{1+} - h_{1+}^*k_2V_{2+} - 2h_{2+}^*k_1V_{1+};$$

$$E_3 = 2k_1V_{1+}k_2\left[V_{2+} - m_1^*(y_1)\right].$$

The presented equation are also valid for the y_{2i} values, which correspond to the training dataset y_{n-}.

According to the equation (7), it follows that there is uncertainty in choosing the sign in front of the square root in the formula for estimation of $\sigma^*(y_1)$. The conditions for choosing the sign before the square root were determined based on the results of statistical modeling of algorithms (7). It should be noted that when implementing the data processing algorithm, we obtain two estimates of the standard deviation of the noise component, characterized by the Gaussian PDF of its values: $\sigma^*(y_1)$ after processing the counts from the sample y_{n+} and $\sigma^*(y_2)$ after processing the counts from the sample y_{n-}. The final estimate of the standard deviation sigma(y) $\sigma^*(y)$ for the PDF (2) is obtained as the arithmetic mean of these estimates, i.e.

$$\sigma^*(y) = \frac{\sigma^*(y_1) + \sigma^*(y_2)}{2}. \tag{8}$$

For the parameter estimation algorithm (7), it should be noted that figure 2 shows a scheme of additional processing of information regarding the values of parameters $\sigma^*(y_1), p_{1+}, U^*_+$, obtained as a result of the data processing algorithm implementation. The data processing algorithm initially calculates estimates for $\sigma^*_{(+)}(y_1), p^*_{(+)1+}, U^*_{(+)+}$, when the "+" sign appears before the square root in (7), and estimates for $\sigma^*_{(-)}(y_1), p^*_{(-)1+}, U^*_{(-)+}$, when the " $-$" sign precedes the square root in (7).

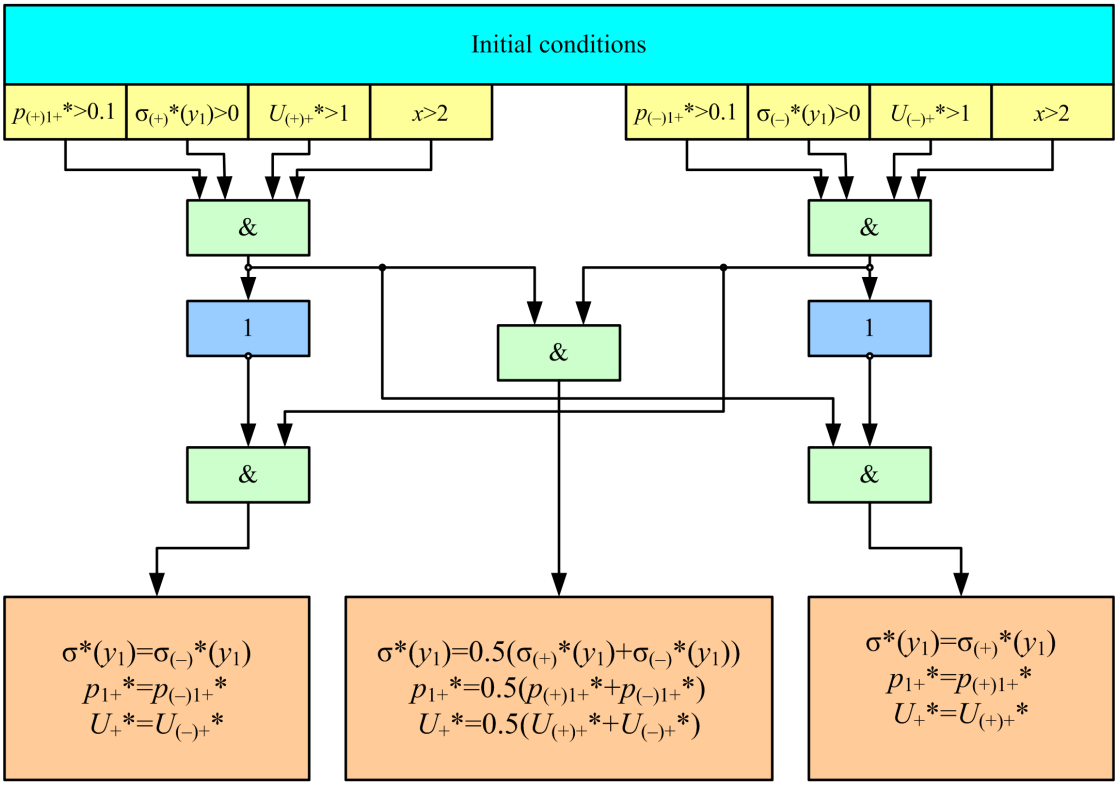


Figure 2: The scheme of additional processing of information regarding the values of parameters.

Thus, when forming the desired estimates of the five parameters of PDF (2) $\sigma(y), U_{-}^{*}, U_{+}^{*}, p_{1}^{*}, p_{2}^{*}$, the original training dataset y_m is divided into two subsets: y_{m+} and y_{m-}. Then, based on data processing from the y_{m+} subset, taking into account the information processing scheme (figure 2), we estimate a portion of the desired parameters U_{+}^{*}, p_{1}^{*}, as well as estimate $\sigma^{*}(y_1)$. Based on data processing from the y_{m-} subset, according to the information processing scheme (figure 2), we estimate a portion of the desired parameters U_{-}^{*}, p_{2}^{*}, as well as estimate $\sigma^{*}(y_1)$. Then, considering eqution (8), we determine the final estimate of $\sigma^{*}(y)$. The additional information processing scheme (figure 2) is constructed using known rules of algebraic logic. We assume that if the $p_{(+)1+}^{*} > 0.1, \sigma_{(+)}^{*}(y_1) > 0, U_{(+)+}^{*} > 1, x = U_{(+)+}^{*}/\sigma_{(+)}^{*}(y_1) > 2$ are simultaneously fulfilled, this corresponds to the situation of forming the logical "one". Similarly, if the conditions $p_{(-)1+}^{*} > 0.1, \sigma_{(-)}^{*}(y_1) > 0, U_{(-)+}^{*} > 1, x = U_{(-)+}^{*}/\sigma_{(-)}^{*}(y_1) > 2$ are simultaneously fulfilled, this also corresponds to the situation of forming the logical "one".

4. Analysis of data processing method for estimating the parameters of multimodal distributions

The analysis of methods for estimating parameters is carried out on the basis of finding the statistical characteristics of the estimates. The most complete one is the probability density function of estimate. However, finding it causes significant difficulties when solving the analysis problem analytically.

In this research, the problem of analyzing the method for estimating five parameters of the PDF was solved on the basis of statistical simulation.

Statistical simulation was performed for 1000 iterations. During the simulation, the following values of the parameters of the estimation algorithm were selected:

$$V_{1+} = 1; V_{2+} = 5; V_{1-} = -1; V_{2-} = -5; k_1 = \tfrac{1}{3}; a_1 = 1.5; k_2 = \tfrac{1}{6}; a_2 = 3.$$

The simulation process was carried out in the following sequence:

1. Generating three independent normal random variables in accordance with PDF (2).
2. Formation of uniform random variable in the interval [0; 1] and comparing it with thresholds in accordance with the values p_1 and p_2.
3. Obtaining a dataset of the noise component.
4. Checking the conditions according to figure 2.
5. Dividing the dataset into two datasets.
6. Estimation of unknown PDF parameters according to formulas (7).
7. Finding the expected values and standard deviations of PDF parameter estimates.

The results of statistical simulation are presented in table 1.

Table 1 contains six options of initial parameters of PDF (2). The numerical values of parameters are shown in table 2.

In general, the simulation results indicate the efficiency of the proposed method for estimating the parameters of the PDF of the noise component. The advantage feature is the low value of the estimates bias and the low level of the standard deviation. To increase the efficiency of

Table 1
The results of statistical simulation.

Initial parameters	Expected value / standard deviation of PDF parameter estimate				
Estimated parameters	$\sigma^*(y)$	U_+^*	U_-^*	p_1^*	p_2^*
Option 1	0.94 / 0.15	4.75 / 0.81	4.22 / 0.79	0.28 / 0.069	0.24 / 0.069
Option 2	0.99 / 0.12	4.81 / 0.42	4.48 / 0.54	0.27 / 0.051	0.26 / 0.047
Option 3	0.81 / 0.14	6.21 / 0.72	5.44 / 1.12	0.23 / 0.064	0.24 / 0.066
Option 4	0.83 / 0.14	6.38 / 0.67	5.39 / 1.07	0.21 / 0.057	0.25 / 0.039
Option 5	0.83 / 0.106	2.64 / 0.24	4.19 / 0.99	0.34 / 0.067	0.21 / 0.073
Option 6	0.84 / 0.085	2.72 / 0.17	4.02 / 0.49	0.32 / 0.049	0.21 / 0.042

Table 2
The numerical values of parameters.

Option number	$\sigma(y)$	U_+	U_-	p_1	p_2	m
1	1	5	4	0.25	0.25	50
2	1	5	4	0.25	0.25	100
3	1	7	5	0.2	0.25	50
4	1	7	5	0.2	0.25	100
5	1	3	4	0.3	0.2	50
6	1	3	4	0.3	0.2	100

estimation more accurate techniques of approximation can be used, for example considered in [28].

5. Conclusions

The paper is devoted to the problem of synthesis and analysis of the method for estimating the parameters of multimodal distribution. Such distributions are often used to describe non-Gausian noise. The paper considers the specific example of interference in the form of chaotic pulsed interference with positive and negative polarity values. Such noise case can be described by a five-parameter PDF according to Tukey's model as a weighted sum of three normal distributions.

The synthesis of the method for estimating the five parameters of the PDF was carried out based on the use of the method of moments and the method of quantiles, which made it possible to obtain the system of equations containing the estimation parameters. The numerical solution of the equations was made possible by approximating the probability integral using the linear function.

The analysis of the method for estimating the five parameters of the PDF was carried out on the basis of statistical simulation. The simulation results have showed satisfactory estimation results.

References

[1] J. Duda, A. Gąsior, Industry 4.0. A Glocal Perspective, Routledge, New York, NY, 2022.

[2] J.-C. André, Industry 4.0: Paradoxes and Conflicts, Wiley, Hoboken, NJ, 2019.

[3] N. S. Kuzmenko, I. V. Ostroumov, Performance Analysis of Positioning System by Navigational Aids in Three Dimensional Space, in: 2018 IEEE First International Conference on System Analysis and Intelligent Computing (SAIC), 2018, pp. 1–4. doi:10.1109/SAIC.2018.8516790.

[4] O. A. Sushchenko, Robust Control of Platforms with Instrumentation, in: 2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON), 2019, pp. 518–521. doi:10.1109/UKRCON.2019.8879969.

[5] J. Lu, Z. Yan, J. Han, G. Zhang, Data-Driven Decision-Making (D3M): Framework, Methodology, and Directions, IEEE Transactions on Emerging Topics in Computational Intelligence 3 (2019) 286–296. doi:10.1109/TETCI.2019.2915813.

[6] J. T. McClave, T. Sincich, Statistics, 13 ed., Pearson, London, 2020.

[7] A. Renyi, Probability Theory, Dover Publications, New York, NY, 2007.

[8] I. V. Ostroumov, N. S. Kuzmenko, Accuracy estimation of alternative positioning in navigation, in: 2016 4th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC), 2016, pp. 291–294. doi:10.1109/MSNMC.2016.7783164.

[9] L. Feng, W. Pingbo, T. Suofu, C. Zhiming, Parameters Estimation for Colored Non-Gaussian Background in Signal Detection, in: 2010 Second International Conference on Computer Modeling and Simulation, volume 2, 2010, pp. 300–303. doi:10.1109/ICCMS.2010.319.

[10] O. A. Sushchenko, Y. M. Bezkorovainyi, V. O. Golytsin, Modelling of Microelectromechanical Inertial Sensors, in: 2019 IEEE 15th International Conference on the Experience of Designing and Application of CAD Systems (CADSM), 2019, pp. 1–5. doi:10.1109/CADSM.2019.8779286.

[11] A. Anand, M. Ram, System Reliability Management: Solutions and Techniques, CRC Press, Boca Raton, 2021.

[12] M. Modarres, K. Groth, Reliability and Risk Analysis, CRC Press, Boca Raton, 2023.

[13] I. Ostroumov, N. Kuzmenko, Risk Assessment of Mid-air Collision Based on Positioning Performance by Navigational Aids, in: 2020 IEEE 6th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC), 2020, pp. 34–37. doi:10.1109/MSNMC50359.2020.9255506.

[14] O. Shmatko, V. Volosyuk, S. Zhyla, V. Pavlikov, N. Ruzhentsev, E. Tserne, A. Popov, I. Ostroumov, N. Kuzmenko, K. Dergachov, O. Sushchenko, Y. Averyanova, M. Zaliskyi, O. Solomentsev, O. Havrylenko, B. Kuznetsov, T. Nikitina, Synthesis of the optimal algorithm and structure of contactless optical device for estimating the parameters of statistically uneven surfaces, Radioelectronic and Computer Systems (2021) 199–213. doi:10.32620/REKS.2021.4.16.

[15] O. Solomentsev, M. Zaliskyi, Correlated Failures Analysis in Navigation System, in: 2018 IEEE 5th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC), 2018, pp. 41–44. doi:10.1109/MSNMC.2018.8576306.

[16] D. J. Smith, Reliability, Maintainability and Risk. Practical Methods for Engineers, 1- ed., Elsevier, London, 2021.

[17] M. Popela, O. Šimon, S. Vaněk, Universal Measuring Device for Antenna Parameter Testing and Radio Traffic Analysis, in: 2022 32nd International Conference Radioelektronika (RADIOELEKTRONIKA), 2022, pp. 01–05. doi:10.1109/RADIOELEKTRONIKA54537.2022.9764958.

[18] O. V. Solomentsev, V. H. Melkumyan, M. Y. Zaliskyi, M. M. Asanov, UAV operation system designing, in: 2015 IEEE International Conference Actual Problems of Unmanned Aerial Vehicles Developments (APUAVD), 2015, pp. 95–98. doi:10.1109/APUAVD.2015.7346570.

[19] H. W. Ott, Electromagnetic Compatibility Engineering, Wiley, New York, NY, 2009.

[20] D. Ma, D.-l. Su, Research on data mining processing methods for electromagnetic environment monitoring results, in: 2010 Asia-Pacific International Symposium on Electromagnetic Compatibility, 2010, pp. 1626–1629. doi:10.1109/APEMC.2010.5475487.

[21] V. P. Kharchenko, N. S. Kuzmenko, I. V. Ostroumov, Identification of unmanned aerial vehicle flight situation, in: 2017 IEEE 4th International Conference Actual Problems of Unmanned Aerial Vehicles Developments (APUAVD), 2017, pp. 116–120. doi:10.1109/APUAVD.2017.8308789.

[22] H. W. Ott, Noise Reduction Techniques in Electronic Systems, Wiley, New York, NY, 1988.

[23] R. S. Odarchenko, S. O. Gnatyuk, T. O. Zhmurko, O. P. Tkalich, Improved method of routing in UAV network, in: 2015 IEEE International Conference Actual Problems of Unmanned Aerial Vehicles Developments (APUAVD), 2015, pp. 294–297. doi:10.1109/APUAVD.2015.7346624.

[24] M. P. Deisenroth, A. A. Faisal, C. S. Ong, Mathematics for Machine Learning, Cambridge University Press, 2020.

[25] I. Prokopenko, Nonparametric Change Point Detection Algorithms in the Monitoring Data, in: Z. Hu, S. Petoukhov, I. Dychka, M. He (Eds.), Advances in Computer Science for Engineering and Education IV, Springer International Publishing, Cham, 2021, pp. 347–360. doi:10.1007/978-3-030-80472-5_29.

[26] M. K. Srivastava, A. H. Khan, N. Srivastava, Statistical Inference: Theory of Estimation, PHI, Delhi, 2014.

[27] H. L. van Trees, Detection, Estimation, and Modulation Theory, Wiley, New York, NY, 2001.

[28] J. Al-Azzeh, A. Mesleh, M. Zaliskyi, R. Odarchenko, V. Kuzmin, A Method of Accuracy Increment Using Segmented Regression, Algorithms 15 (2022) 378. doi:10.3390/a15100378.

Application of artificial intelligence in digital marketing

Ihor V. Ponomarenko[1], Volodymyr M. Pavlenko[2], Oksana B. Morhulets[2], Dmytro V. Ponomarenko[3] and Nataliia M. Ukhnal[4]

[1]State University of Trade and Economics, 19 Kyoto Str., Kyiv, 02156, Ukraine
[2]Kyiv National University of Technologies and Design, 2 Mala Shyianovska Str., Kyiv, 01011, Ukraine
[3]International University of Business and Law, 9 Heroiv Ukrainy Str., Mykolaiv, 54007, Ukraine
[4]Academy of Financial Management, 38 Druzhby Narodiv Blvd., Kyiv, 01014, Ukraine

Abstract

Identification of the main directions of using artificial intelligence to optimize the marketing strategies of companies in the digital environment is important in the conditions of intensifying competition on the Internet. Artificial intelligence is considered as a tool for qualitative transformations in the use of digital marketing tools based on various information generated in the global network. The methodological basis of the study is a comprehensive analysis of scientific approaches to the practice of implementing artificial intelligence in the field of digital marketing, the formation of an information base for modeling, and the identification of optimal machine learning algorithms to ensure the competitiveness of brands on the Internet. A scheme of the main sources of information, which must be used by the company for the implementation of artificial intelligence algorithms in the process of increasing the effectiveness of digital marketing tools use, has been developed. Digital marketing tools are presented, which are used to establish communications with the target audience in the long term and ensure an economically feasible level of conversion. The main stages of companies' interaction with the audience on the Internet using modern machine learning algorithms are presented. The main directions of using artificial intelligence in digital marketing have been characterized, which enable the company to achieve a high level of loyalty among users based on personalized interaction models.

Keywords

artificial intelligence, big data, content, digital marketing, machine learning, optimization, target audience

CS&SE@SW 2023: 6th Workshop for Young Scientists in Computer Science & Software Engineering, February 2, 2024, Kryvyi Rih, Ukraine

i.v.ponomarenko.stat@gmail.com (I. V. Ponomarenko); pavlenko.vm@knutd.edu.ua (V. M. Pavlenko); morgulets.ob@knutd.edu.ua (O. B. Morhulets); schumi7@ukr.net (D. V. Ponomarenko); ukhnalnm@gmail.com (N. M. Ukhnal)
https://knute.edu.ua/blog/read/?pid=41329&en (I. V. Ponomarenko);
https://en.knutd.edu.ua/university/faculties/ftecs/ (V. M. Pavlenko);
https://en.knutd.edu.ua/university/faculties/fcci/ktgrb/ (O. B. Morhulets);
https://www.linkedin.com/in/dmytro-ponomarenko7 (D. V. Ponomarenko);
https://www.linkedin.com/in/nataliia-ukhnal-207720273 (N. M. Ukhnal)
0000-0003-3532-8332 (I. V. Ponomarenko); 0000-0003-2163-8508 (V. M. Pavlenko); 0000-0003-4985-0359 (O. B. Morhulets); 0009-0002-2904-3904 (D. V. Ponomarenko); 0000-0002-8562-9355 (N. M. Ukhnal)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

CEUR-WS.org/Vol-3662/paper22.pdf
CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

1. Introduction

Digitization processes force companies to pay significant attention to interaction with users on the Internet. The role of digital communications is gradually increasing, as the result of demographic processes is the replacement of older generations by representatives of more innovatively oriented consumers. Generation Y has certain characteristics of conservative behavior, but they are prone to relatively active use of innovative technologies in everyday life. Along with this, representatives of the Z and Alpha generations belong to the digitized generation, as they were born during the period of intensive development of the Internet and the active introduction of various gadgets to the market [1].

Digitization leads to the transformation of the behavior of consumer groups and the growth of their dependence on innovative technologies. A huge number of modern users spend a significant part of their time on the Internet every day for work, study, leisure, etc. Digital technologies significantly simplify the performance of various tasks and the search for relevant information. Accordingly, modern generations choose more innovative models of behavior and consumption, which leads to the transformation of various types of economic activity. Companies to ensure a sufficient level of competitiveness in the markets of operation on an ongoing basis integrate advanced approaches and technologies into their activities [2, 3].

The process of interaction with users involves the development and implementation of marketing strategies that allow companies to promote products on the market and ensure an economically justified level of profitability. By using effective digital marketing tools, companies get the opportunity to identify their target audience and establish close, long-term relationships with users. The development of technology leads to the evolution of digital marketing and the emergence of more effective tools that help increase the conversion rate. Social communications migrated from the offline to the online environment, acquiring specific characteristics of interaction between users and companies. The orientation of a significant number of modern users to communication in the digital environment stimulates the development of various social networks, which are characterized by certain differences in the construction of communications and the demonstration of thematic content. There are leaders in the social media market, along with this, innovation and a high level of competition stimulate the launch of new products. In 2023, the social network and microblogging service Twitter started rebranding to X, which involves not only changing the brand name but also bringing the existing services and functionality of this network in line with the realities of the modern market. In 2021, Facebook was rebranded as Meta, due to the need to create a virtual reality universe that would function as a social media for digital user interaction [4].

The functioning of companies in a digital environment and the use of modern marketing tools enable companies to accumulate large volumes of various information. It is advisable to use specialized web analytics services for data collection. Along with this, the search for relevant information can be carried out thanks to the use of various methods that have gained significant distribution in the field of Data science. Effective methods of collecting big data in real time include site parsing, which allows for generating relevant information on legal grounds. The information obtained from various sources acts as a valuable resource for finding directions for optimizing the company's marketing strategy in the digital environment and achieving economically feasible results in specific time intervals. Automated real-time data collection

allows companies to quickly identify existing risks and make relevant effective decisions, which is impossible to achieve when using traditional statistical methods of information collection [5].

The development of the market of cloud services has led to the appearance on the market of specialized companies that allow based on powerful servers to accumulate and process large volumes of various information [6]. The presented technology has led to the active development and introduction of various machine learning algorithms used to identify hidden relationships in accumulated data. Artificial intelligence, based on machine learning algorithms and characterized by the ability to learn by the changing influence of internal and external environmental factors, is very popular in the modern world [7, 8, 9].

The purpose of this work is to study the peculiarities of the accumulation of big data and its processing thanks to artificial intelligence to increase the efficiency of digital marketing tools used. The paper considers the main algorithms of machine learning, which are implemented within the framework of artificial intelligence. The integration of artificial intelligence into digital marketing tools will allow to increase in the level of personalized models accuracy of interaction with customers and will contribute to increasing the level of the target audience loyalty.

2. Related works

The functioning of modern companies in the digital environment and the presence of significant competition requires the search for innovative approaches that will allow them to achieve an economically justified level of conversion due to the loyalty of the target audience. Thanks to the use of modern mathematical algorithms by scientifically based methodological approaches for processing heterogeneous information, it is possible to optimize the use of resources available in the company. The development of server technologies has made it possible to implement effective machine learning algorithms that allow the processing of big data and quickly provide results for adjusting marketing strategies. A comprehensive analysis of research shows that there is a significant interest among scientists in identifying new directions for the use of artificial intelligence in the field of digital marketing.

A comprehensive analysis of the features of artificial intelligence use in the field of marketing in modern conditions was carried out by Dumitriu and Popescu [10]. The authors emphasize the key role of digitization processes as a locomotive for the development of the global economic environment, all types of economic activity, and individual companies. A four-stage model is presented, which allows for an increase in the visibility of the company's web resources in the digital environment based on artificial intelligence algorithms. The necessity of using machine learning in the process of improving the system of search engine optimization and identification in an automated mode of high-frequency keywords is proven. Updating the list of keywords makes it possible to ensure a high level of communication with the target audience by the preferences and behavior patterns of users in search engines.

The study of content's role in social media in building effective models of interaction with the target audience is presented by Shahbaznezhad et al. [11]. The selection of relevant content and the formation of an effective content plan make it possible to attract the attention of a large number of users and keep their attention for a long time. By driving interest in their social

media pages, companies have the opportunity to promote products and increase conversion rates. The article by Banerjee [12] is devoted to the features of content selection for social media thanks to artificial intelligence. The authors consider the issues of identifying fake content and building trust with subscribers based only on reliable information.

Balaji et al. [13] proved the importance of using social media for interaction with users, which is connected with the desire of modern generations to actively interact in the digital environment. The process of communications leads to the generation of big data on an ongoing basis, which makes it possible to accumulate valuable information for improving marketing strategies promptly respond to changes in user behavior, and satisfy identified needs by companies. The authors present the most effective machine learning algorithms, which are advisable to use for processing data generated in social media and developing effective management solutions based on the obtained results. Thanks to the use of artificial intelligence, companies get the opportunity to increase the number of followers on social media and generate high interest in their products.

3. Models and methods

The implementation of various machine learning algorithms involves the use of large data that is accumulated by the company in the digital environment and can be used for modeling and optimization of existing processes. Figure 1 presents the main sources of information that a company can use when integrating artificial intelligence into digital marketing tools.

According to the presented approach, it is advisable to collect data in three main directions, which is explained by the peculiarities of the company's interaction with other Internet participants. To collect data from the company's own websites and social media, it is advisable to use specialized web analytics tools that connect to the company's resources and collect information about various activities on an ongoing basis. Web analytics is also used to evaluate the activity of advertising campaigns, but the traffic for advertising messages comes to the company's Internet resources [16]. Accordingly, thanks to web analytics, the effectiveness of the implementation of advertising measures is also evaluated on the company's resources. When setting up web analytics tools by science-based approaches, the system of metrics used for data collection is determined. The flexibility of this approach allows companies to update the indicators used at any time, adapting to changes in the influence of internal and external environmental factors. In some cases, it is possible to use web analytics tools to research competitors' web resources, but this approach allows companies to get only a limited set of data [17].

To collect socio-economic and demographic indicators on the Internet, including data on functioning markets, main competitors, and consumers, it is advisable to use publicly available sources. First of all, it is possible to use information from international organizations, national statistical organizations, and state administration bodies [18]. Along with this, it is possible to download data from the sites of non-governmental organizations and thematic communities. The presented information is mainly in an aggregated form and is intended for public use, as it is not a commercial secret. However, the application of these data makes it possible to adapt the marketing strategy in the offline and online environment to the existing realities, which helps to increase the competitiveness of the company in the long term.

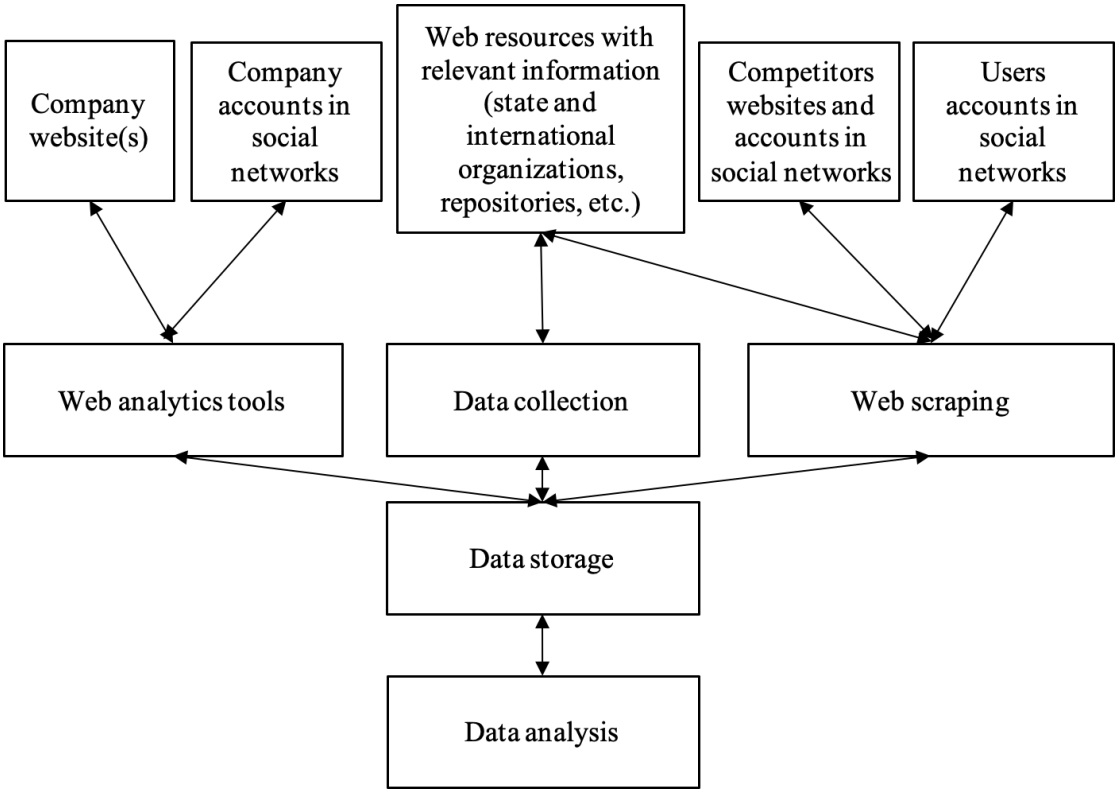


Figure 1: The main sources of information for artificial intelligence applications in digital marketing [14, 15].

In the conditions of digitalization, the scraping of web resources becomes an important tool for gathering information, as it allows companies to automate the search and accumulation of relevant data. In most cases, scrapers allow to quickly accumulate data that can be collected by company employees when browsing competitors' web resources, but people need much more time to search and collect valuable information [19]. For marketing purposes in the digital environment, scrapers may collect the following information about competitors and the market environment: product prices, textual content, related competitor information, product reviews and ratings, socio-demographic characteristics of customers and visitors, popular hashtags, available promotions and discounts, keywords, e-mails and other personal data of users, photos, videos, and other media content. In certain cases, unethical or illegal use of scraping is observed, which leads to obtaining personal data that is not publicly available [20].

There are a large number of tools that are characterized by certain differences and are used to interact with the target audience and establish communications in certain conditions. An effective marketing strategy in the digital environment involves the simultaneous use of several tools, the combination of which varies depending on the specifics of the influence of internal and external environmental factors. Comprehensive influence on the target audience thanks to the use of selected digital marketing tools allows to achieve the highest possible level of conversion and provides prerequisites for a loyal attitude of users to the brand over a long period. It should

be noted that the set of tools may change with the transformation of the company's marketing strategy in the digital environment. Figure 2 shows the main digital marketing tools.

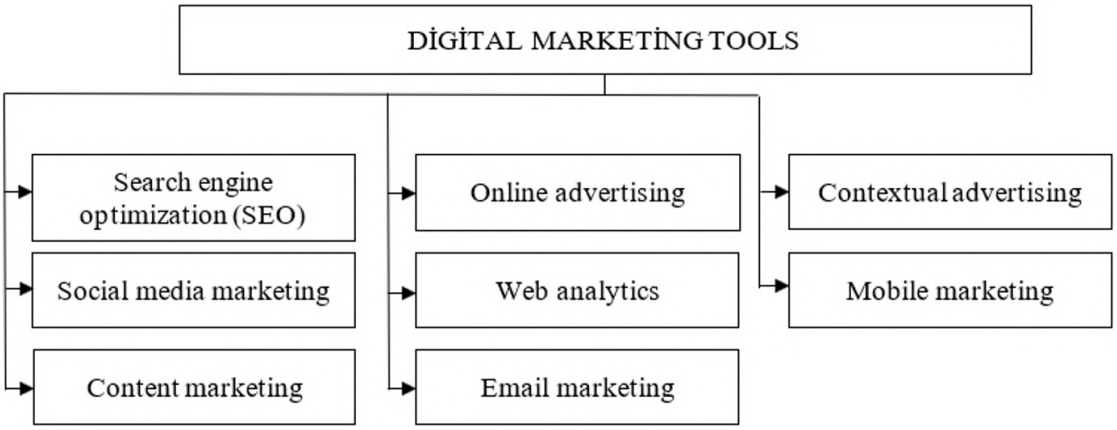


Figure 2: Digital marketing tools [21].

Information exchange during the implementation of digital marketing tools has a reciprocal nature, because thanks to the use of existing tools, the company gets the opportunity to collect complex data about related processes. Along with this, based on the received big data, a comprehensive analysis of the investigated phenomena is carried out, including the implementation of machine learning algorithms, and the obtained results are the basis for optimizing the use of digital marketing tools. Accumulating up-to-date information on an ongoing basis makes it possible to adjust the marketing strategy in the digital environment and achieve effective results in long-term periods [22].

Users act as an important source of information for the company, interacting with the brand through web resources (including social media, advertising messages, and other digital communication channels). The use of web analytics tools and other approaches to gathering information allows the company to accumulate large data that is used in the implementation of machine learning algorithms. Figure 3 shows companies' interaction with the audience on the Internet.

In the process of the company's interaction with users, various web resources are used, which is explained by the expediency of using a certain number of digital communication channels. Depending on the personalities of the target audience and the specifics of the company's activities and product characteristics, various interaction channels can be used. However, for communications with the target audience, in many cases, the company's official website, specialized landing pages with promotions or individual products, various social media, Internet advertising, etc. are used. The use of a certain number of digital marketing channels allows a brand to increase the reach of the target audience and ensure an economically justified level of conversion.

When users interact with specific web resources of the company, web analytics tools collect information from the selected metrics system. Accumulated information is transferred to servers and processed using machine learning algorithms. Following the scientific methodology, the data

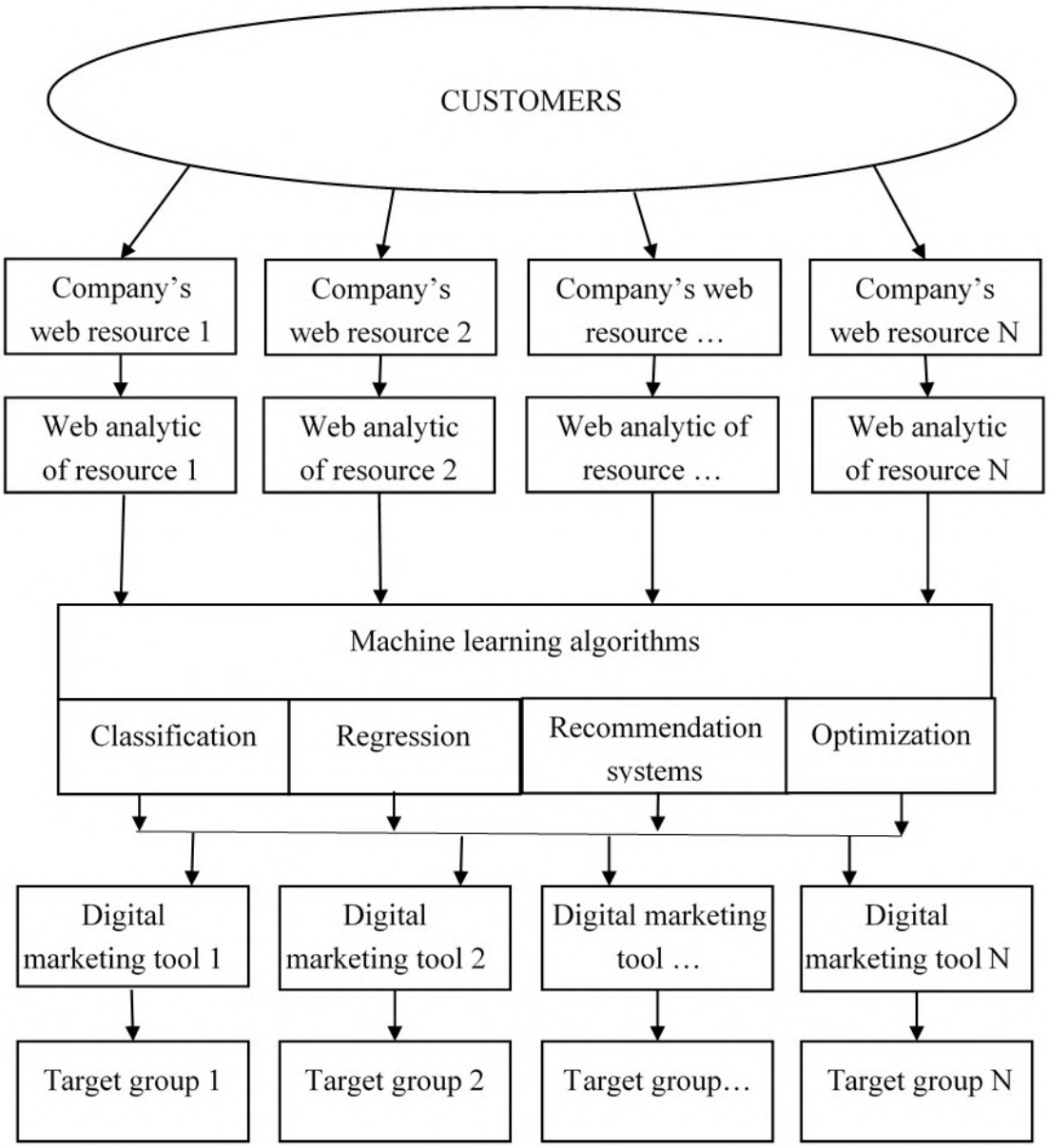


Figure 3: Companies' interaction with the audience on the Internet [23, 24, 25].

processing system is adjusted, which includes the selection of the machine learning algorithm. Based on the data received from the user, a specific mathematical model is implemented and the optimal digital marketing tool is selected for further interaction with the relevant client.

Choosing a model of interaction with specific consumers based on complex calculations allows a company to achieve effective results with a high level of probability. Due to user identification, relevant content and optimal communication channels are selected. The interaction of the company with the user in the digital environment by the identified consumer behavior and

psychological characteristics is positively perceived by the client and leads to the formation of a loyal relationship with a specific brand over a long period [26, 27].

In modern conditions, machine learning algorithms are used as elements of artificial intelligence to obtain optimal results. When using artificial intelligence in digital marketing based on big data, there is a constant process of improving the realization of relevant machine learning algorithms and obtaining more accurate results. The increase in accuracy is achieved due to the self-learning of models by the action of internal and external environmental factors with a constant search for optimal solutions.

The use of artificial intelligence in the improvement of the company's marketing strategy in the digital environment allows companies to obtain a set of advantages based on the use of comprehensive information on the studied phenomena. The application of various thematic content significantly expands the possibilities for the application of machine learning algorithms. It should be noted that various groups of users communicate with the brand and other participants through the use of various behavioral models. The social networks are configured to use specific content in the communication process. The collection of heterogeneous information and its processing thanks to the use of machine learning algorithms allows for a more detailed investigation of existing processes and the identification of hidden relationships [28].

Given the importance of using artificial intelligence in digital marketing, it is necessary to describe the main areas of integration of this approach to optimize interaction between companies and the target audience. First of all, it is advisable to pay attention to the following areas of application:

1. *Analysis of big data.* Machine learning algorithms allow companies to process large amounts of information and identify hidden relationships between the company, its products, and the target audience. Thanks to different approaches, photo, audio, and video content are transformed into digital form and used for comprehensive analysis. Along with this, cause-and-effect models are implemented to determine the influencing factors on consumer behavior and forecasts are made regarding the trends in the development of phenomena related to the marketing digital environment. Artificial intelligence gradually adapts to existing circumstances and allows you to build dynamic regression and predictive models online [29].
2. *Personalization of content.* Integration into artificial intelligence of various classification models, including clustering by a large number of indicators, allows for dividing the population of users into specific groups. The identification of hidden relationships leads to the identification of groups of customers with special needs, which involves the implementation of specialized communication models to satisfy individual consumers. For each of the groups, specialized content is selected, which with a high level of probability will be suitable for the presented target audience. Thanks to the use of artificial intelligence, an individual consumer will not only receive relevant content but will also perceive interaction with the company as a personalized approach. Formation in the mind of the client of an individual approach on the part of the company leads to the construction of close long-term communications.
3. *Content generation.* Modern artificial intelligence allows companies not only to process a variety of data but also to generate a variety of content based on input information.

OpenAI company developed ChatGPT, which is very popular in today's world. Among the applied areas, it is advisable to pay attention to digital marketing, because using this service it is possible to generate relevant content by a text request. The generated text information can be used for posting on the company's web resources, writing scripts for advertisements, etc. Along with this, the received text should be used to communicate with the target audience on social media. Interaction with users and answering topical questions in social networks should be prompt, which involves the use of relevant textual content. ChatGPT allows companies to optimize the marketing of social networks and ensure a high level of interest and loyalty in the target audience. The market of artificial intelligence is actively developing, which led to the appearance on the market of Copilot (Microsoft), Gemini (Google), Bedrock (Amazon), Llama 2 (Meta), etc. Along with this, OpenAI has developed an innovative product based on Dall-E 3 and ChatGPT, which allows generating complex images based on text description. The resulting visual content contains several drawn objects that can interact with each other [30].

4. *Customer support.* The presented direction of using artificial intelligence in digital marketing combines to a certain extent the two previous directions, as it allows interaction with users, identifying their needs based on requests, and providing reliable answers with a high level of probability. The development of mathematical algorithms makes it possible to endow chatbots with certain human traits, which are positively perceived by the target audience. Thanks to the evolution and improvement of artificial intelligence, chatbots get the opportunity not only to classify a request and provide an answer from the existing library of sentences but also to independently generate answers. Along with text assistants, voice services are actively developing, and are gaining popularity among modern consumers [31].
5. *Sentiment analysis.* Companies need to receive objective information about the attitude of the target audience to brands and corresponding products. Along with the basic information provided by web analytics services about visiting resources on the Internet and revealing interest in products through views and purchases, it is also advisable to conduct a detailed analysis of user reactions in comments. When evaluating the relationship of the target audience, it is possible to use likes and other buttons with reactions, along with this, users like to leave various comments. Thanks to sentiment analysis based on artificial intelligence, it is possible to identify user points of view based on comments, emoticons, and other graphic content. Identifying the relationship of the target audience to certain activities of the brand in social media allows c to adjust the company's strategy in the digital environment to achieve an optimal result [32].

4. Further research

The obtained results show the prospects of using artificial intelligence to optimize the use of digital marketing tools by companies. The development of server technologies and programming languages makes it possible to identify new, more productive machine learning algorithms. In parallel, specialized programming languages are developing, first of all, it is necessary to pay attention to Python and libraries for implementing the corresponding mathematical algorithms.

These directions are important for the development of science as a whole and the development of innovative methodological approaches in the field of digital marketing. Deepening the study of certain machine learning algorithms according to the needs of specific digital marketing tools will significantly increase the effectiveness of interaction with the target audience, focusing on the implementation of a personalized approach and the generation of unique content according to the requests of an individual client. Studying issues related to the creation of a personal experience through artificial intelligence in the process of interaction between the brand and the client will contribute to the growth of ties between the participants of the communication process over a long period.

5. Conclusions

AI-based digital marketing tools are becoming a necessary component of companies' strategies on the Internet, as they allow to ensure the necessary level of competitiveness. Thanks to the application of big data generated continuously in the digital environment, machine learning algorithms are self-learning and constantly find more effective ways to improve marketing strategies. The use of artificial intelligence is constantly expanding due to the discovery of new directions and opportunities in digital marketing. The introduction of innovative approaches will be carried out in the future continuously due to significant competition between companies and the formation of a stable demand for advanced information products that allow identifying the target audience and ensuring a high level of loyalty. Along with this, in the process of implementing marketing strategies in the digital environment, there is a constant need for relevant content to form close communications with consumers.

References

[1] S. Azimi, Y. Andonova, C. Schewe, Closer together or further apart? Values of hero generations Y and Z during crisis, Young Consumers 23 (2022) 179–196. doi:10.1108/YC-03-2021-1300.

[2] Y. Li, J. Dai, L. Cui, The impact of digital technologies on economic and environmental performance in the context of industry 4.0: A moderated mediation model, International Journal of Production Economics 229 (2020) 107777. doi:10.1016/j.ijpe.2020.107777.

[3] O. B. Morgulets, T. M. Derkach, Information and communication technologies managing the quality of educational activities of a university, Information Technologies and Learning Tools 71 (2019) 295–304. doi:10.33407/itlt.v71i3.2831.

[4] P. Fernandez, Facebook, Meta, the metaverse and libraries, Library Hi Tech News 39 (2022) 1–5. doi:10.1108/LHTN-03-2022-0037.

[5] M. Y. S. Bak, J. B. Plavnick, A. D. Dueñas, M. T. Brodhead, S. M. Avendaño, A. J. Wawrzonek, E. Weber, S. N. Dodson, N. Oteto, The use of automated data collection in applied behavior analytic research: A systematic review, Behavior Analysis: Research and Practice 21 (2021) 376–405. URL: https://www.researchgate.net/publication/355152581. doi:10.1037/bar0000228.

[6] V. Oleksiuk, O. Oleksiuk, The practice of developing the academic cloud using the Proxmox

VE platform, Educational Technology Quarterly 2021 (2021) 605–616. doi:10.55056/etq.36.
[7] L. O. Fadieieva, Enhancing adaptive learning with Moodle's machine learning, Educational Dimension 5 (2021) 1–7. doi:10.31812/ed.625.
[8] I. A. Pilkevych, D. L. Fedorchuk, M. P. Romanchuk, O. M. Naumchak, Approach to the fake news detection using the graph neural networks, Journal of Edge Computing 2 (2023) 24–36. doi:10.55056/jec.592.
[9] O. V. Klochko, V. M. Fedorets, V. I. Klochko, Empirical comparison of clustering and classification methods for detecting internet addiction, CTE Workshop Proceedings 11 (2024) 273–302. doi:10.55056/cte.664.
[10] D. Dumitriu, M. A.-M. Popescu, Artificial Intelligence Solutions for Digital Marketing, Procedia Manufacturing 46 (2020) 630–636. doi:10.1016/j.promfg.2020.03.090.
[11] H. Shahbaznezhad, R. Dolan, M. Rashidirad, The Role of Social Media Content Format and Platform in Users' Engagement Behavior, Journal of Interactive Marketing 53 (2021) 47–65. doi:10.1016/j.intmar.2020.05.001.
[12] T. J. Banerjee, A System of Content Analysis of Social Media using AI and NLP, International Journal of Research in Engineering, Science and Management 4 (2021) 132–136. URL: https://journal.ijresm.com/index.php/ijresm/article/view/844.
[13] T. K. Balaji, C. S. R. Annavarapu, A. Bablani, Machine learning algorithms for social media analysis: A survey, Computer Science Review 40 (2021) 100395. doi:10.1016/j.cosrev.2021.100395.
[14] U. Sivarajah, Z. Irani, S. Gupta, K. Mahroof, Role of big data and social media analytics for business to business sustainability: A participatory web context, Industrial Marketing Management 86 (2020) 163–179. doi:10.1016/j.indmarman.2019.04.005.
[15] G. Barbera, L. Araujo, S. Fernandes, The Value of Web Data Scraping: An Application to TripAdvisor, Big Data and Cognitive Computing 7 (2023) 121. doi:10.3390/bdcc7030121.
[16] J. Maintz, F. Zaumseil, Tracking content marketing performance using web analytics: tools, metrics, and data privacy implications, International Journal of Internet Marketing and Advertising 13 (2019) 170–182. URL: https://ideas.repec.org/a/ids/ijimad/v13y2019i2p170-182.html.
[17] C. Mahfoudh, B. Othmane, The Role of Web Analytics in Online Marketing, in: S. Sedkaoui, M. Khelfaoui, R. Benaichouba, K. Mohammed Belkebir (Eds.), International Conference on Managing Business Through Web Analytics, Springer International Publishing, Cham, 2022, pp. 411–423. doi:10.1007/978-3-031-06971-0_29.
[18] O. Plaksiuk, O. Yakushev, O. Yakusheva, L. Moisieienko, Analysis and Assessment of Human Capital in the Regions of Slovakia, Economics. Ecology. Socium 7 (2023) 13–25. doi:10.31520/2616-7107/2023.7.3-2.
[19] M. A. Khder, Web Scraping or Web Crawling: State of Art, Techniques, Approaches and Application, International Journal of Advances in Soft Computing & Its Applications 13 (2021) 144–168. doi:10.15849/IJASCA.211128.11.
[20] J. Boegershausen, H. Datta, A. Borah, A. T. Stephen, Fields of Gold: Scraping Web Data for Marketing Insights, Journal of Marketing 86 (2022) 1–20. doi:10.1177/00222429221100750.
[21] M. K. Peter, M. Dalla Vecchia, The Digital Marketing Toolkit: A Literature Review for the

Identification of Digital Marketing Channels and Platforms, in: R. Dornberger (Ed.), New Trends in Business Information Systems and Technology: Digital Innovation and Digital Business Transformation, Springer International Publishing, Cham, 2021, pp. 251–265. doi:10.1007/978-3-030-48332-6_17.

[22] C. Katsikeas, L. Leonidou, A. Zeriti, Revisiting international marketing strategy in a digital era: Opportunities, challenges, and research directions, International Marketing Review 37 (2020) 405–424. doi:10.1108/IMR-02-2019-0080.

[23] M. Broersma, Audience Engagement, in: The International Encyclopedia of Journalism Studies, John Wiley & Sons, Ltd, 2019, pp. 1–6. doi:10.1002/9781118841570.iejs0060.

[24] S. N. Samsudeen, M. Kaldeen, Impact of digital marketing on purchase intention, International Journal of Advanced Science and Technology 29 (2020) 1113–1120. URL: https://www.researchgate.net/publication/341670094.

[25] R. Motoryn, K. Prykhodko, B. Ślusarczyk, P. Żegleń, Evaluation of regional features of electronic commerce in Europe, Statistical Journal of the IAOS 38 (2022) 1339–1347. URL: https://content.iospress.com/articles/statistical-journal-of-the-iaos/sji220938.

[26] A. Aluri, B. S. Price, N. H. McIntyre, Using Machine Learning To Cocreate Value Through Dynamic Customer Engagement In A Brand Loyalty Program, Journal of Hospitality & Tourism Research 43 (2019) 78–100. doi:10.1177/1096348017753.

[27] I. Ponomarenko, I. Panasiuk, V. Pavlenko, O. Panasiuk, O. Kalmykov, Use of Neural Networks for Pattern Recognition in E-Commerce, CEUR Workshop Proceedings 3179 (2021) 407–415. URL: https://ceur-ws.org/Vol-3179/Short_15.pdf.

[28] T. Ribeiro, J. L. Reis, Artificial Intelligence Applied to Digital Marketing, in: Á. Rocha, H. Adeli, L. P. Reis, S. Costanzo, I. Orovic, F. Moreira (Eds.), Trends and Innovations in Information Systems and Technologies, Springer International Publishing, Cham, 2020, pp. 158–169. doi:10.1007/978-3-030-45691-7_15.

[29] M. S. Mazorchuk, T. S. Vakulenko, A. O. Bychko, O. H. Kuzminska, O. V. Prokhorov, Cloud technologies and learning analytics: web application for PISA results analysis and visualization, CTE Workshop Proceedings 8 (2021) 484–494. doi:10.55056/cte.302.

[30] S. C. Gamoura, H. İ. Koruca, K. B. Urgancı, Exploring the Transition from "Contextual AI" to "Generative AI" in Management: Cases of ChatGPT and DALL-E 2, in: Z. Şen, Ö. Uygun, C. Erden (Eds.), Advances in Intelligent Manufacturing and Service System Informatics, Springer Nature Singapore, Singapore, 2024, pp. 368–381. doi:10.1007/978-981-99-6062-0_34.

[31] M. A. Camilleri, C. Troise, Live support by chatbots with artificial intelligence: A future research agenda, Service Business 17 (2023) 61–80. doi:10.1007/s11628-022-00513-9.

[32] N. V. Babu, E. G. M. Kanaga, Sentiment Analysis in Social Media Data for Depression Detection Using Artificial Intelligence: A Review, SN Computer Science 3 (2021) 74. doi:10.1007/s42979-021-00958-1.

Modern methods of energy consumption optimization in FPGA-based heterogeneous HPC systems

Oleksandr V. Hryshchuk, Sergiy P. Zagorodnyuk

Taras Shevchenko National University of Kyiv, 64/13 Volodymyrska Str., Kyiv, 01601, Ukraine

Abstract

High-Performance Computing (HPC) systems play a pivotal role in addressing complex computational challenges across various domains, but their escalating energy consumption has raised concerns regarding sustainability and operational costs. This paper presents a comprehensive investigation into the parametrization and modeling of energy consumption in heterogeneous HPC systems, aiming to provide valuable insights for optimizing energy efficiency while preserving performance. We begin by characterizing the heterogeneity within modern HPC environments, which encompass diverse hardware components, such as CPUs, GPUs, FPGAs, and accelerators. Our research delves into modeling techniques, leveraging heuristics methods and statistical approaches to construct accurate predictive models for energy consumption. Furthermore, we explore the integration of dynamic power management strategies, such as DVFS (Dynamic Voltage and Frequency Scaling) and task scheduling, to optimize energy usage without compromising performance. This paper provides a vital foundation for sustainable HPC practices, enabling researchers and practitioners to make informed decisions for achieving enhanced energy efficiency without sacrificing computational performance.

Keywords

high-performance computing (HPC), FPGA, power modeling, power analysis, heterogeneous computing, power saving, task scheduling,

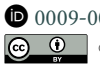
1. Introduction

Today's large-scale computing systems, such as data centers and high-performance computing clusters (HPCs), are severely limited by power and cooling costs for extremely large-scale (or exascale) problems. The steady increase in electricity consumption is a growing concern for several reasons, such as cost, reliability, scalability, and environmental impact. Nowadays data centers use 200 TWh per year and contribute near 0.3% of whole carbon emissions in the world, when entire complex of ICT (Information and computing technology) devices produce up to 2% of it [1]. Best case scenario model predicts that in 2030 ICT will share 8% of whole electricity consumption in the world [2], while worst case scenario anticipate 51% of global electricity usage. This potential increase in power consumption and, sequentially, cost of computing operations leads researcher and engineers to investigate and develop new techniques and approaches to optimize power management in HPC systems and in ICD domain in general.

CS&SE@SW 2023: 6th Workshop for Young Scientists in Computer Science & Software Engineering, February 2, 2024, Kryvyi Rih, Ukraine

oleksandr_hryshchuk@knu.ua (O. V. Hryshchuk); szagorodniuk@gmail.com (S. P. Zagorodnyuk)

0009-0007-9926-4231 (O. V. Hryshchuk); 0000-0003-3415-7746 (S. P. Zagorodnyuk)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

CEUR-WS.org/Vol-3662/paper17.pdf

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

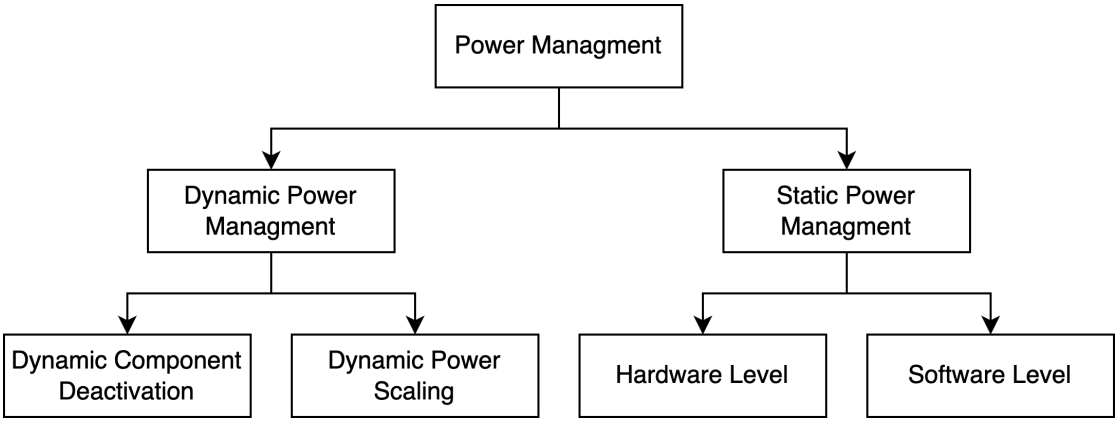


Figure 1: General classification of power management methods in computer systems.

Present-day there are set of methods and approaches to resolve this energy optimization issue, mainly only for homogeneous CPU-based HPC systems. General taxonomy of this techniques, suggested in [3] and depicted on figure 1 and can be divided into two main groups SPM (static power management) and DPM (dynamic power management). SPM methods, divided in two separate groups (for hardware and software level management) usually defined during design time and cannot be changed in runtime. Hardware SPM techniques can be detailed and split into three separate groups [3]:

1. Circuit level
2. Logic level
3. Architecture level

DPM methods widely used in HPC [4] systems can be divided into two main groups – DCD (Dynamic component Deactivation), based on predictive and heuristic approaches, and DPS (Dynamic Power Scaling), like resource throttling and DVFS (Dynamic Voltage Frequency Scaling). This techniques can be a foundation for more complicated optimization methods, in example, task scheduling based on DVFS [5] or DCD heuristics applications [4].

Methods described before can be used on different hardware platforms, both homogeneous (well-studied nowadays) and heterogeneous (with GPU, TPU, FPGA and CGRA), which became popular in HPC according to a survey on Deep Learning hardware accelerators for heterogeneous HPC Platforms [6]. At the same time number of scientific papers on energy-aware optimization for HPC systems with FPGA controllers are extremely low (1-3 per year), compared to all researches about “FPGA heterogeneous computing” (see figure 2 with data obtained from app.dimensions.ai) which indicates a limited number of solutions in this domain, so this work will be focused on heterogeneous applications of energy-aware optimizations in HPC systems.

2. Energy optimization theory

2.1. Optimization problem definition for task scheduling

In introduction section was mentioned that optimization techniques can be divided into hardware and software types, first of them are case-specific for different variations of hardware like CPU, memory chips, NIC, etc., while software-defined approaches can be generalized and provide a solution for disparate equipment with same characteristics/types, in example, homogeneous or heterogeneous GPU and TPU-based HPC clusters [7]. Such software solutions are often leads to energy-efficient task-scheduling methods, optimization problem for which can be defined in a way that described next.

For a finite set of jobs(task) J and a finite set of resources $R, time(j, r)$ is a function, that returns time of execution of job $j \in J$ on resource $r \in R$ [4].Then scheduling can be described as task of finding a set of start times $\{s_1, s_2, \ldots, s_{|J|}\}$ for jobs, allocated to resources $\{a_1, a_2, \ldots, a_{|J|}\}$ in conditions where:

$$\forall s_x : \nexists s_y : s_x \leq s_y + \text{time}\,(y, A_y) \wedge s_y \leq s_x + \text{time}\,(x, A_x) \wedge a_x = a_y, \forall a_x : x \in R \quad (1)$$

Additional optimization conditions (see equation 2) can be applied to provided scheduling, where optimization criteria can be finding maximum or minimum, depending on formulation of a function which involves simple metrics such as execution time, consumed energy, etc. [4].

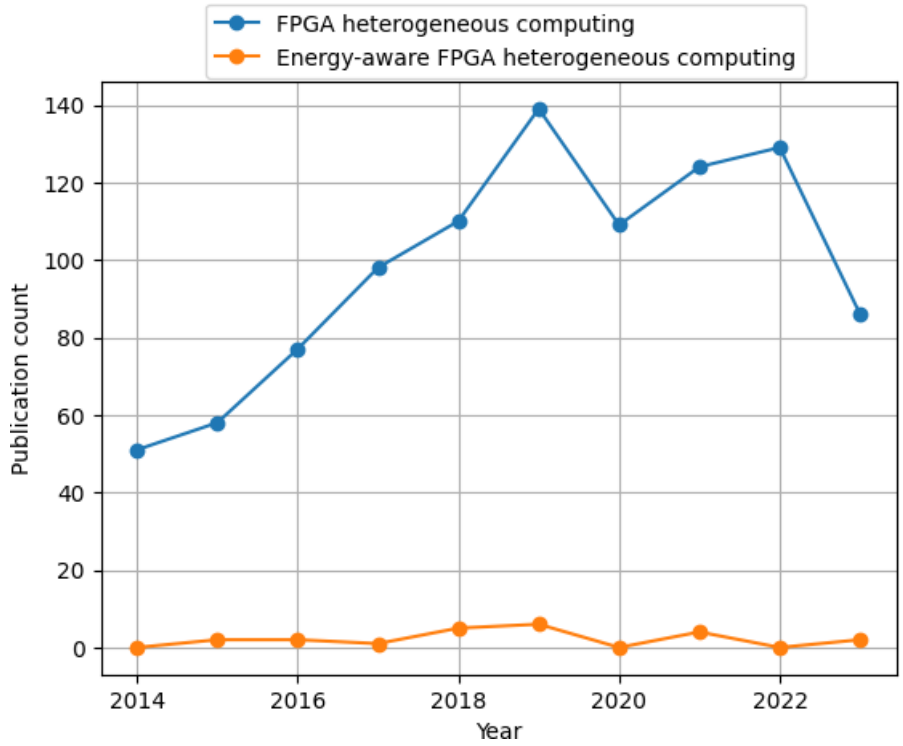


Figure 2: Count of scientific publications per year on topic "FPGA heterogeneous computing" and "Energy-aware FPGA heterogeneous computing" from 2014 to 2023.

$$\min / \max \left(\text{ OptimizationCriteria } \left(\left\{ s_1, s_2, \ldots, s_{|J|} \right\}, \left\{ a_1, a_2, \ldots, a_{|J|} \right\} \right) \right) \tag{2}$$

This model is extremely simplified and does not suitable for real applications due to several reasons – it assumes that one resource can take only one task at the time, number of available resources always equal or higher than number of jobs to complete and does not include impact of communication between tasks on nodes or computing elements. To resolve these problems and adapt model to real world upgraded model was suggested [4] – for two tasks x and y from set of jobs pairs D, P_j is set of devices, which can be assigned for job $j \in J$, time of communication between jobs obtained from function $comm(x, y, a_x, a_y)$, then solution is a set of assignments A_j and start times $\{s_1, s_2, \ldots, s_{|J|}\}$ for each job, like it described in equations 3-6:

$$\forall x \in A_j : x \in P_j \tag{3}$$

$$\forall s_x : \nexists s_y : s_x \leq s_y + \text{time}\,(y, A_y) \wedge s_y \leq s_x + \text{time}\,(x, A_x) \wedge A_x \cap A_y = \emptyset \tag{4}$$

$$\forall \{x, y\} \in D : s_x + \text{time}\,(x, A_x) + \text{comm}\,(x, y, A_x, A_y) \leq s_y \tag{5}$$

With optimization condition:

$$\min / \max \left(\text{ OptimizationCriteria } \left(\left\{ s_1, s_2, \ldots, s_{|J|} \right\}, A_1, \ldots, A_{|J|}, D \right) \right) \tag{6}$$

This method involves enumeration of all jobs for all available resources, which leads to idea that solution can not be found in polynomial time, and it was proved that problem of energy-efficient active time [8] scheduling is NP-Complete [5], so to be able use this model there can be a two possible ways – use predefined constraints and precalculated configurations or use heuristic methods, in example genetic algorithms [9], to find solution during runtime.

2.2. Optimization criteria

General optimization problem was described in previous section, and to be used in real HPC systems in requires properly defined optimization criteria. Existing solutions in this domain based on energy consumption metric (EC), or can take under consideration other properties, in example, execution time, etc. [4]. Power consumption can be described via energy itself (in joules or watts), or can be represented with more complicated models like instruction per joule or power per watt [10]. This approach used in Green500 rating as FLOPS per Watt metric [11].

More sophisticated can use combination of following metrics such as EC (energy consumption), ExecT (execution time), utilization, average weighted time, wait time, power, Pareto front, AST, AFT, clock frequency, work(job) per energy, reliability, electricity cost, temperature, EDP, EDF, Number of cores, Probability of execution, branch transition rate, cache efficiency, issue width [4]. In example new algorithm was proposed for reformed scheduling method with energy consumption constraint (RSMECC), based on AST, AFT and energy consumption metrics [12]. This algorithm can make it possible to more efficiently solve a wide range of computing tasks, including in the field of neural networks, complex 3D modeling and artificial intelligence.

3. Cluster architecture

Nowadays HPC clusters widespread around the world in different forms and variations, but generally main part of them are based on homogeneous massive parallel processor architecture (MPP), which inherited from older NUMA (non-uniform memory access) architecture [13]. This approach looks similar to shared-memory technology, but in this case each processor in cluster is connected to it's own part of memory and create entity of single independent node, which connected with other nodes via network interface card and common network (see figure 3). Absence of shared memory between nodes (not including common NAS) simplifies design and reduces inefficient components therefor improving scalability and stability of HPC system [13]. At the same time due to lack of shared memory, a processor core in one group must employ a different method to exchange data and coordinate with cores of other processor groups [14]. This issue become more visible for heterogeneous systems, based on CPUs form different series or types, or even for GRID computing systems [15].

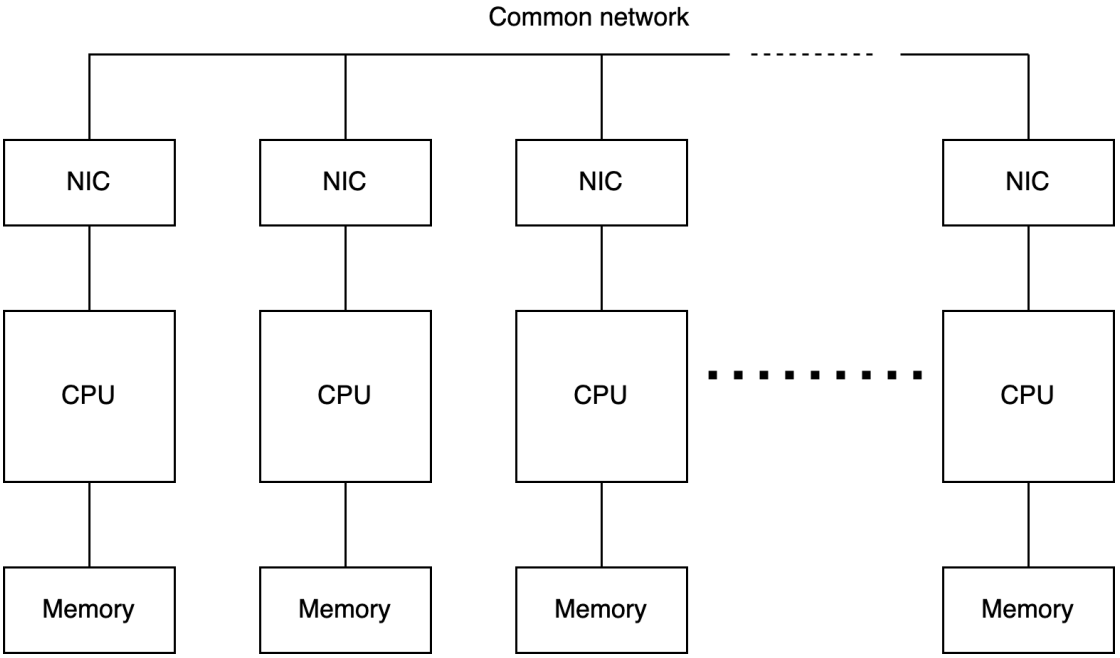


Figure 3: MPP HPC cluster architecture.

Another popular approach for building HPC systems is usage of symmetric multi-processors (SMP). It embodies a category of parallel architectures that harness the power of multiple processor cores to enhance performance by leveraging parallel processing, all the while upholding a unified memory structure that spans the entirety of the parallel computing system [13].

An SMP defines a self-contained and self-sustaining computer system equipped with all the subsystems and components essential for fulfilling the demands and facil-itating the execution of various applications. It can operate independently to support user applications designed as shared-memory multi-threaded programs, serve as one among several equivalent subsystems

in a scalable MPP systems or commodity clus-ter, and work as a throughput computer for the simultaneous execution of independent concurrent tasks [14]. General architecture of SMP system depicted on figure 4.

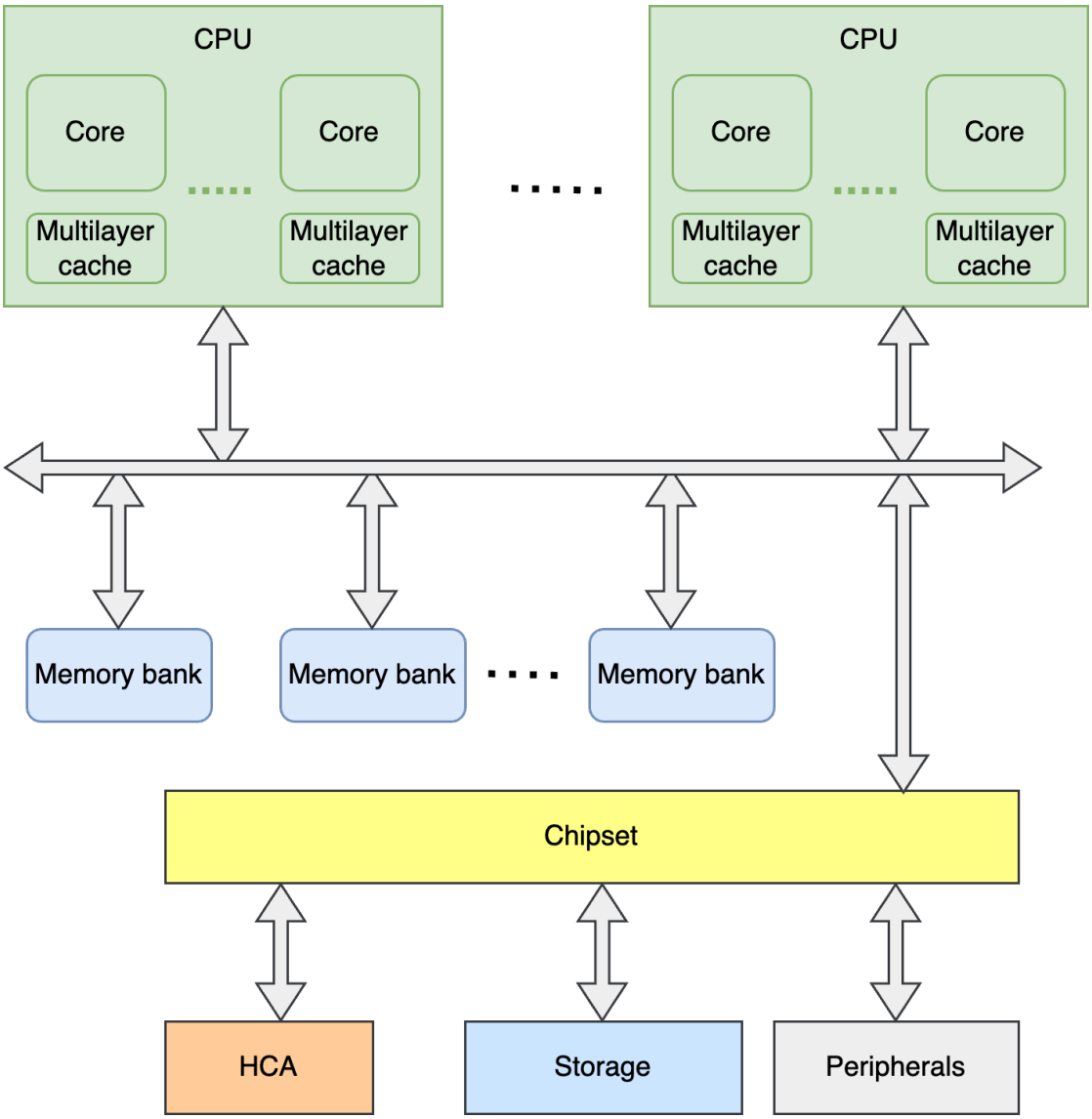


Figure 4: Internal architecture of SMP HPC system.

3.1. Heterogeneous cluster architecture comparison

Heterogeneous computing in HPC refers to the utilization of diverse hardware accelerators, like general purpose graphic processing unit (GPGPU), field programmable gate array (FPGA), coarse-grained reconfigurable array (CGRA) [15] and specialized coprocessors, alongside traditional

CPU. This approach harnesses the strengths of different computing components to optimize performance and energy efficiency, making it particularly well-suited for workloads that can benefit from parallel processing. Most common heterogeneous clusters involve usage of coupled CPU and GPGPU as single node, therefore nowadays exists energy efficient solutions for this kind of HPC system, which was analyzed in [4].

But FPGA in same time in HPC is a new type of accelerators and less studied as it was shown in Introduction section of this paper. But nowadays there are existing works on this topic, in example the technique of cooperative CPU, GPU and FPGA task execution, based on EngineCL framework was suggested in [16]. Also, new approach, called Cooperative Heterogeneous Acceleration with Reconfigurable Multi-devices (CHARM) was proposed for multi hybrid accelerated cluster with GPU and FPGA coupling, which was implemented in “Albireo-nodes” in Cygnus cluster, based on CPU Intel Xeon Gold, GPU NVIDIA Tesla V100 x4 and FPGA Nallatech 520N with Intel Stratix10 [17]. Architecture of this nodes shown of figure 5.

Characteristic comparison for Cygnus supercomputer node and heterogeneous system from EngineCL test setup shown on table 1. At the same time, for EngineCL was shown that performance improvement from heterogeneity was obtained for all benchmark tasks ("Matrix multiplication", "Mersenne Twister", "Watermarking", "Sobel Filter", "Nearest Neighbor", "AES Decrypt"), but energy consumption improvement was detected only for "Sobel Filter" [16], which leaves a research gap for searching energy-optimization methods for this kind of system.

Table 1
Comparison of Cygnus and EngineCL setup node specifications.

Characterisic	Cygnus	EngineCL test setup
CPU	Intel Xeon Gold x2	Intel Core i7-G700k
GPU	Nvidia Tesla V100x4 (32 Gb x4)	Nvidia GeForce GTX Titan X (12 Gb)
FPGA	Intel Stratix 10x2	Altera DE5NET Stratix V
RAM	192GB	64GB
Number of nodes	32 GPU+FPGA, 46 CPU-only	1
Energy-efficiency	N/A	1 of 6 benchmark tasks

Consequently, this two works have a lack of energy consumption optimization for described systems, and despite existing methods of power management and optimization described in survey of FPGA optimization methods for data center energy efficiency [18]. Finding “general” solution for FPGA-kind of system is complicated due to the necessity of reconfiguring of hardware for each specific task (job), but nevertheless, energy optimization constraints with proper criteria, described in “Energy optimization theory” section of this paper can be applied to multi-hybird hardware FPGA systems to optimize power consumption.

4. Conclusions

This paper shows modern theories and approaches for power consumption planning and optimizations for heterogeneous HPC systems, including optimization model for MPP system, described in third section of this paper. As this problem in NP-complete, heuristics approaches

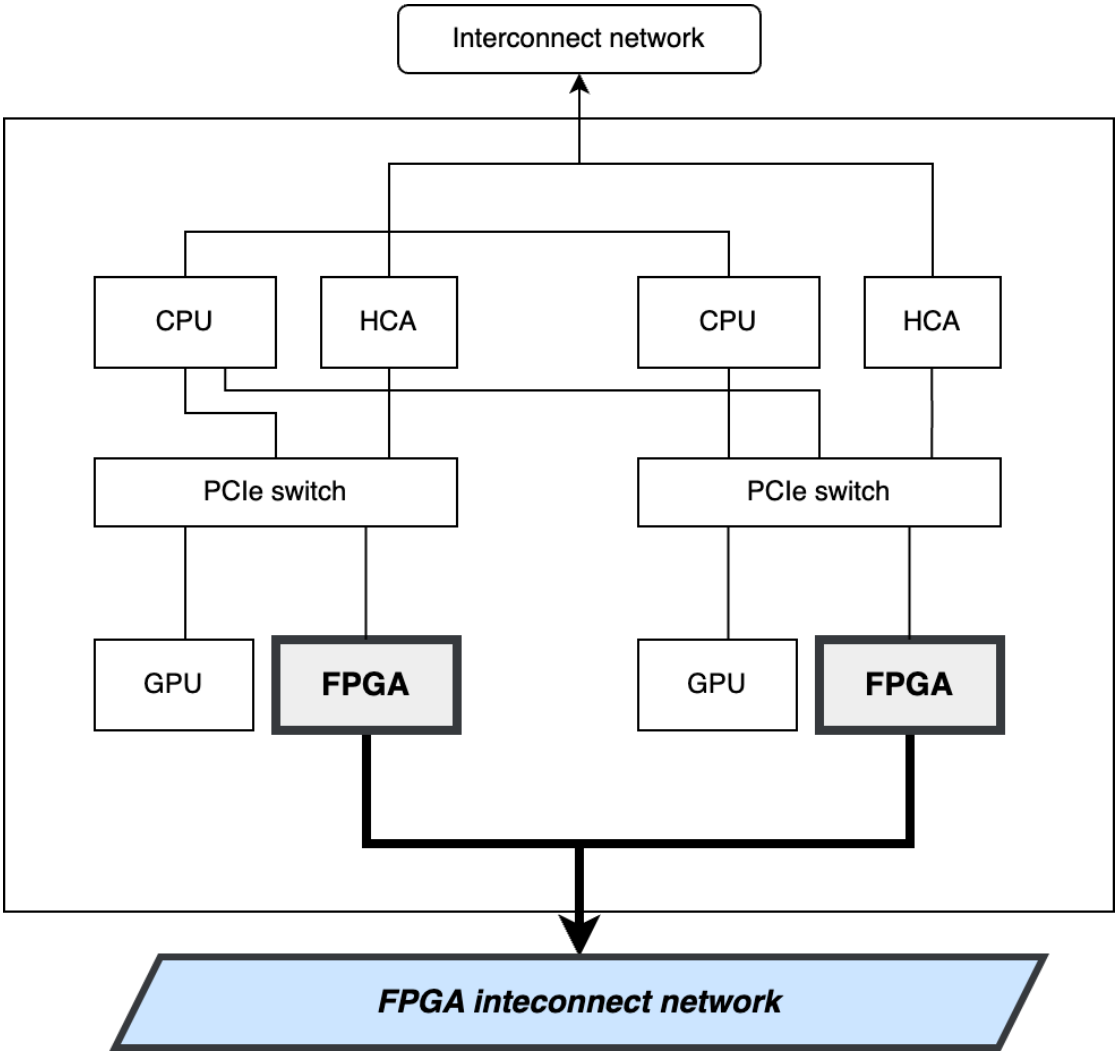


Figure 5: Internal architecture of Albireo-node from Cygnus cluster.

for finding solutions was mentioned. Results from mentioned solutions can be implemented on hardware or software level via DPM technologies. At the same time mentioned solutions is well suited to only CPU-GPU coupled systems, but not for CPU-GPU-FPGA coupled systems. For last one there is existing power management techniques, like easy-to use in FPGA DCD, but the is a lack of schedulers and general approaches for implementing solution from theoretical optimal model. Therefore, future work involves further search ways of amplification methods, including heuristic solutions of power consumption planning in FPGA-coupled HPC systems.

References

[1] N. Jones, How to stop data centres from gobbling up the world's electricity, Nature 561 (2018) 163–166. doi:10.1038/d41586-018-06610-y.

[2] A. S. G. Andrae, T. Edler, On Global Electricity Usage of Communication Technology: Trends to 2030, Challenges 6 (2015) 117–157. doi:10.3390/challe6010117.

[3] J. Haj-Yahya, A. Mendelson, Y. B. Asher, A. Chattopadhyay, Energy Efficient High Performance Processors: Recent Approaches for Designing Green High Performance Computing, Springer, 2018.

[4] B. Kocot, P. Czarnul, J. Proficz, Energy-Aware Scheduling for High-Performance Computing Systems: A Survey, Energies 16 (2023). doi:10.3390/en16020890.

[5] S. Saha, M. Purohit, NP-completeness of the Active Time Scheduling Problem, 2021. URL: http://arxiv.org/abs/2112.03255.

[6] C. Silvano, D. Ielmini, F. Ferrandi, L. Fiorin, S. Curzel, L. Benini, F. Conti, A. Garofalo, C. Zambelli, E. Calore, et. al., A Survey on Deep Learning Hardware Accelerators for Heterogeneous HPC Platforms, 2023. doi:10.48550/arXiv.2306.15552.

[7] V. Raca, S. Umboh, E. Mehofer, B. Scholz, Runtime and energy constrained work scheduling for heterogeneous systems, Journal of Supercomputing 78 (2022) 17150–17177. doi:10.1007/s11227-022-04556-7.

[8] J. Chang, H. N. Gabow, S. Khuller, A Model for Minimizing Active Processor Time, in: L. Epstein, P. Ferragina (Eds.), Algorithms – ESA 2012, Lecture Notes in Computer Science, Springer, 2012, pp. 289–300. doi:10.1007/978-3-642-33090-2_26.

[9] A. Cocaña-Fernández, J. Ranilla, L. Sánchez, Energy-efficient allocation of computing node slots in HPC clusters through parameter learning and hybrid genetic fuzzy system modeling, Journal of Supercomputing 71 (2015) 1163–1174. doi:10.1007/s11227-014-1320-9.

[10] M. Safari, R. Khorsand, Energy-aware scheduling algorithm for time-constrained workflow tasks in DVFS-enabled cloud environment, Simulation Modelling Practice and Theory 87 (2018) 311–326. doi:10.1016/j.simpat.2018.07.006.

[11] T. Scogland, J. Azose, D. Rohr, S. Rivoire, N. Bates, D. Hackenberg, Node variability in large-scale power measurements: perspectives from the Green500, Top500 and EEHPCWG, in: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, SC '15, Association for Computing Machinery, New York, NY, USA, 2015. doi:10.1145/2807591.2807653.

[12] Y. Hu, J. Li, L. He, A reformed task scheduling algorithm for heterogeneous distributed systems with energy consumption constraints, Neural Computing and Applications 32 (2020). doi:10.1007/s00521-019-04415-2.

[13] T. Sterling, M. Brodowicz, M. Anderson, High Performance Computing: Modern Systems and Practices, Morgan Kaufmann, 2017.

[14] S. Ramos, T. Hoefler, Modeling communication in cache-coherent SMP systems: a case-study with Xeon Phi, in: Proceedings of the 22nd international symposium on High-performance parallel and distributed computing, HPDC '13, Association for Computing Machinery, 2018, pp. 97–108. doi:10.1145/2462902.2462916.

[15] P. S. Käsgen, M. Weinhardt, C. Hochberger, A Coarse-Grained Reconfigurable Array for High-Performance Computing Applications, in: 2018 International Conference on Re-

ConFigurable Computing and FPGAs (ReConFig), 2018, pp. 1–4. doi:10.1109/RECONFIG.2018.8641720.

[16] M. Dávila, R. Nozal, R. Gran Tejero, M. Villarroya, D. Suárez Gracia, J. Bosque, Cooperative CPU, GPU, and FPGA heterogeneous execution with EngineCL, The Journal of Supercomputing 75 (2019). doi:10.1007/s11227-019-02768-y.

[17] T. Boku, N. Fujita, R. Kobayashi, O. Tatebe, Cygnus - World First Multihybrid Accelerated Cluster with GPU and FPGA Coupling, in: Workshop Proceedings of the 51st International Conference on Parallel Processing, ICPP Workshops '22, Association for Computing Machinery, 2023, pp. 1–8. doi:10.1145/3547276.3548629.

[18] M. Tibaldi, C. Pilato, A Survey of FPGA Optimization Methods for Data Center Energy Efficiency, IEEE Transactions on Sustainable Computing (2023) 343–362. doi:10.1109/TSUSC.2023.3273852.

Conference platform metadata and functions: existing platforms analysis and ontology-based approach

Yevhenii B. Shapovalov[1,2], Viktor B. Shapovalov[1,2]

[1]Ukrainian Scientific Center for the Development of Information Technologies, 180 Antonovycha Str., Kyiv, 03150, Ukraine

[2]National Center "Junior Academy of Science of Ukraine", 38-44 Dehtiarivska Str., Kyiv, 04119, Ukraine

Abstract

The academic landscape is currently marred by the rise of predatory conferences, which are more concerned with profits than with promoting genuine scholarly discourse. These events often lack the stringent peer review processes necessary to ensure the dissemination of high-quality research. To address this issue, the academic community urgently requires robust conference management platforms capable of discerning and filtering out such dubious gatherings. These platforms should be designed not only to facilitate the search and participation in conferences but also to rigorously evaluate and ensure the quality of the conferences listed. Incorporating insights from various studies, it becomes evident that an ideal platform would merge the functionalities of advanced peer review, decision-making tools based on ethical considerations, and concrete quality metrics. By implementing these features, conference management platforms can become the bulwark against the dilution of research integrity and play a pivotal role in nurturing the value of scientific conferences. The literature underscores a shift towards focusing on the core issue of quality in conferences, suggesting that a multifaceted platform could successfully address both the problems of questionable conference quality and predatory nature.

Keywords

predatory conferences, conference platforms, ontology, metadata, automation

1. Introduction

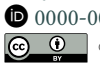
The collective literature points to the growing concern of predatory conferences, which are characterized by their profit-driven motives and lack of rigorous peer review, leading to the dissemination of low-quality research. The term "questionable conference", as suggested by McCrostie [1], aligns with a more precise identification of these events based on specific red flags. This nomenclature shift directs the focus towards the quality problem rather than merely labeling the conferences as predatory.

Chartier's work underlines the rise of such conferences, hinting at the critical need for reliable platforms that ensure trusted selection and peer review processes [2]. Pecorari introduces a tool designed to assist researchers in making informed decisions about conference participation, thereby equipping them with the means to discern quality. This suggests that platforms could integrate such tools to facilitate ethical decision-making for prospective attendees [3].

CS&SE@SW 2023: 6th Workshop for Young Scientists in Computer Science & Software Engineering, February 2, 2024, Kryvyi Rih, Ukraine

sjb@man.gov.ua (Y. B. Shapovalov); svb@man.gov.ua (V. B. Shapovalov)

0000-0003-3732-9486 (Y. B. Shapovalov); 0000-0001-6315-649X (V. B. Shapovalov)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

CEUR-WS.org/Vol-3662/paper20.pdf

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

Ibrahim's discussion on the adverse effects of predatory practices on scientific literature further substantiates the necessity for heightened awareness and educational resources within conference platforms to combat these issues [4].

Stevic's hybrid model provides a framework for evaluating conference quality through measurable aspects. By embedding such a model into conference platforms, it could systematically assess and ensure the quality of conferences listed or hosted on the platform [5].

Martins' proposition of new quality metrics tailored for conferences could be utilized by platforms to standardize quality assessment, moving beyond subjective perceptions to more concrete and comparative measures [6].

Laplante's emphasis on the importance of paper quality for the reputation of conferences suggests that platforms could enforce strict peer review guidelines and quality checks for paper submissions [7].

Berndtsson et al. [8], Hagemann-Wilholt et al. [9], Bentlage et al. [10], Iana et al. [11] explore various aspects of conference systems and platforms, from subjective quality assessment to the practicalities of online conference management and the development of recommendation systems. Integrating these insights could lead to the creation of multifaceted platforms that not only focus on quality control but also enhance the overall user experience through pervasive computing, metadata curation, and personalized recommendations.

By converging these diverse perspectives, the argument shifts from evaluating the quality of conferences to addressing the core issue of quality itself. This approach underscores the potential of conference platforms to serve as a solution to the problems of quality and the predatory nature of some conferences. Such platforms could incorporate advanced peer review processes, ethical decision-making tools, quality metrics, and pervasive computing to ensure the integrity and value of scientific conferences.

Currently, Ukrainian Research Information System (URIS) is being developing [12] and the system that will ensure accounting of the scientific data is developing as sub-system of it [13]. Therefore, to ensure consideration of existing experience during development of national platform of conferences, this study aims to analyze metadata and functions of existing conference platforms.

2. Methods

2.1. Data collection

We analyzed the 6 most well-known conference platforms. The platforms that being analyses included two groups of software: informational-oriented and process-oriented. The first group principle is just to provide information and structure. This group included ConfiDent (https://www.confident-conference.org/index.php/Main_Page), Conference Index (https://conferenceindex.org/), OpenResearch (https://www.openresearch.org/wiki/Main_Page) and WikiCFP (http://www.wikicfp.com/cfp/). In other hand, process-oriented systems provide whole cycle of processes that are being required to providing conferences. It could include processes of submission, reviewing and publishing. As examples we took Morressier and EasyChair.

2.2. Data processing

We provided data aggregation using Pivot tables in MS Excel. To ensure correct structural data representation based on the type of data that we collected, we added colonum that describes the type of characteristics of data that we collected. This colonum included Presence of event class, Basic metadata of conferences, User fields, Identifiers, Functionality, Planned functionality, Organizational data, and Additional data. It is worth noting that we analyze web pages and scholar papers, but it ensures that we obtained full information about the systems due to some fields of functions may be not described or shown through the privacy settings. We have legginged in all systems where it was possible to avoid to, but there is still a possibility of it. That's why we each function or field in our analysis could be in the state "no data" for cases when we did not find the relative function /field or "yes" for cases when this field or fucntion was available. Those classes were used as filters in Pivot table and graphs were built for each such class separately.

3. Results

3.1. Description ot the conference platform

3.1.1. Description of the informational-oriented systems

ConfIDent [9], Conference Index, OpenResearch [9], and WikiCFP [11] are four distinct platforms, each with its unique focus on academic and scientific events, albeit with some overlap in their offerings.

ConfIDent is a platform that aims to ensure the persistent accessibility of scientific events such as conferences in high quality. It appears to focus on the longevity and quality of the information regarding scientific meetings, ensuring that records of these events are maintained over time 1.

Conference Index categorizes and indexes conferences globally. It organizes information by category (such as engineering, physics, health science), by tag (like education, environment, medicine), and by country, highlighting the international span of its listings. The platform emphasizes its role in reaching potential participants through its indexing services 2.

OpenResearch seeks to make descriptive metadata on conferences and other scientific events permanently accessible, with high quality through automated processes and scientific data curation. It serves a broad audience including researchers, universities, specialized societies, and funding agencies. The platform provides information on thousands of events and event series, and allows browsing by fields of science, type of content, and region, thus facilitating a targeted search for scientific events and resources.

WikiCFP is a semantic wiki dedicated to Calls for Papers (CFPs) in the fields of science and technology. It hosts over 100,000 CFPs and is utilized by a large number of researchers monthly. The platform is essentially a repository and dissemination point for upcoming calls related to conferences, workshops, and journals, focusing on the early stages of conference participation and paper submission. This table provides a high-level overview of the platforms' purposes and services, although it does not capture all requested parameters due to limited information

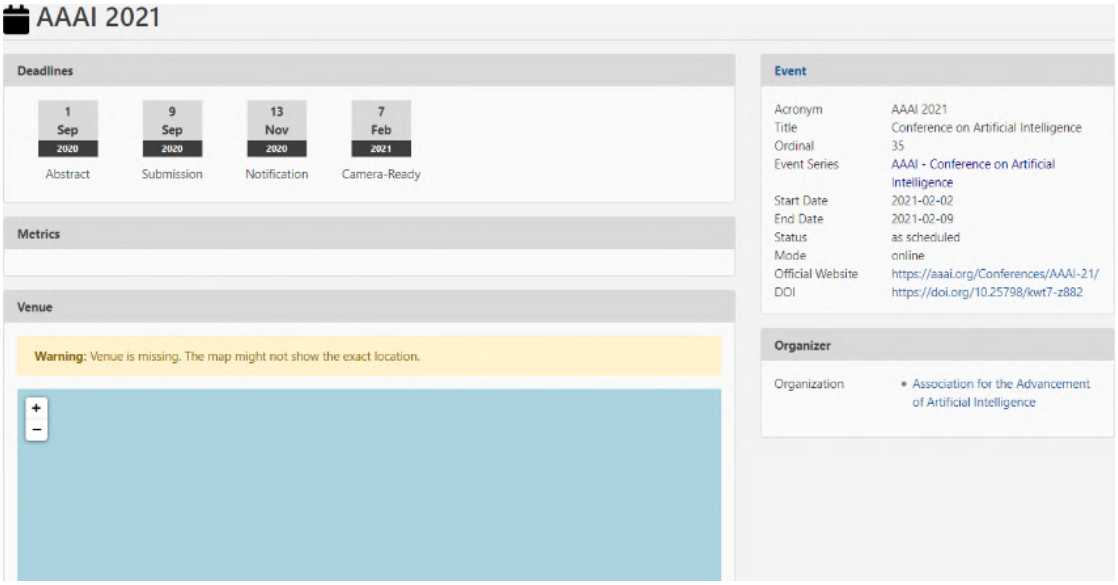


Figure 1: General view of ConfIDent conference platform.

availability. For specific details such as the number of visitors or more precise ownership and location data, further in-depth research or direct inquiry with the platforms may be necessary.

A general description of those systems is presented in table 1.

3.1.2. Description of the process-oriented systems

Morressier and EasyChair are both platforms that cater to the needs of the academic and research communities, albeit with different approaches and services.

Morressier is designed to support the pre-publishing journey of scholarly communications. It provides an array of services that support the early stages of research dissemination, from managing hybrid and virtual conferences to facilitating journal submissions, peer-review workflows, and implementing AI-powered integrity checks. The platform's vision is to enhance efficiency and trust in the scientific communication process using technology.

EasyChair offers a comprehensive conference management system that supports the organization of scientific conferences, both virtually and in-person. It encompasses a virtual conference solution to facilitate scientific conferences of any size, a conference management system for handling everything from program committees to publishing proceedings, and a registration system that allows for the creation of complex registration forms and online payment processing in multiple currencies. EasyChair also includes 'Smart Slides', a feature that enables the publication and distribution of presentation slides, and 'Smart CFP', a tool for publishing conference calls for submissions. In addition, their publishing services offer a seamless submission-to-publication process for reviewed content in various scientific areas.

These platforms highlight the diverse technological tools available to modern researchers and conference organizers, aiming to streamline the scientific communication lifecycle from

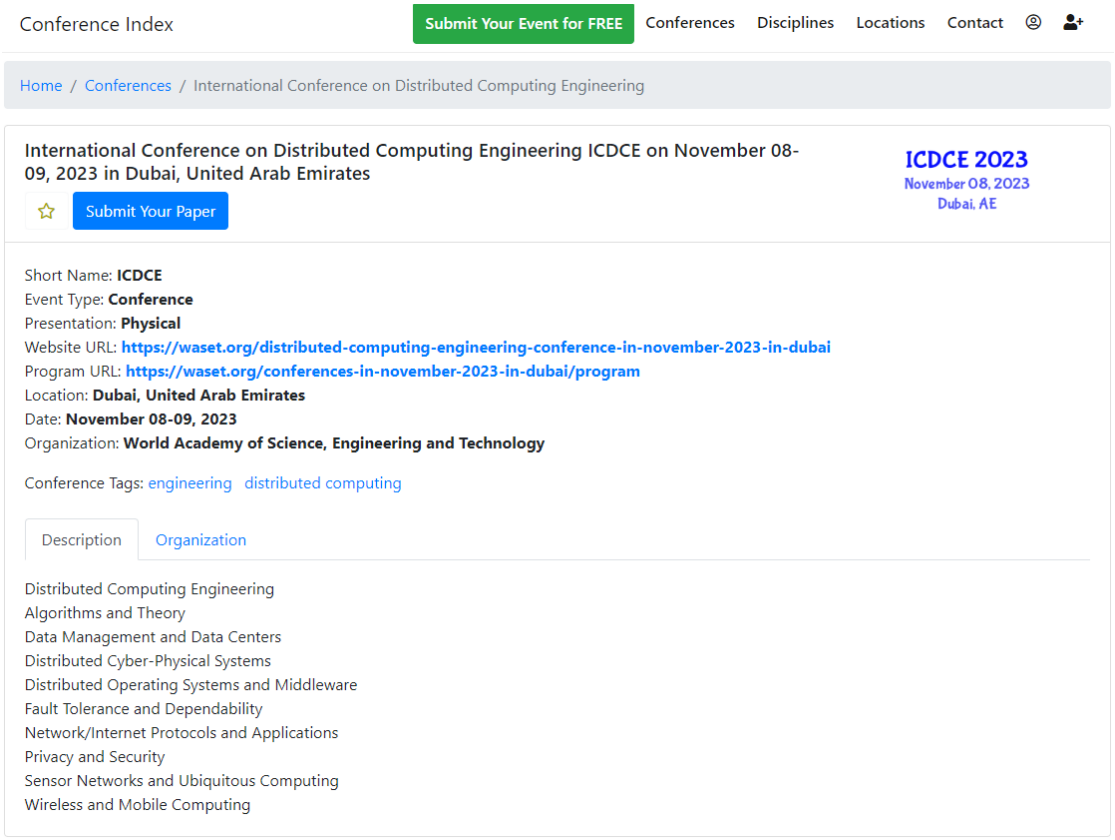


Figure 2: General view of Conference Index conference platform.

initial call for papers to post-conference publication and dissemination. A general description of those systems is presented in table 2.

3.2. Frequency of fields usage

Fields such as 'Title', 'Type', 'Country', 'City' used in four systems, indicate that users often delineate conferences based on these categorical distinctions. These could serve as primary filters when searching for events relevant to an individual's location preference or subject matter.

The fields that are used in 3 systems are 'Venue', 'The term of notification of assignment', 'The start date', 'The official website of the conference', 'Serial number (series of events)', 'Organizers', 'Final submission', 'Direction', 'Date of completion', 'Acronym', and 'A series of events' — reflect a middle tier of frequency in usage. Those fields underscore general aspects of conferences including the logistical aspects of conferences, including information on event timelines, locations, and organizational details, which are critical for planning and attendance. The inclusion of 'Final submission' indicates a focus on the deadlines for academic contributions, while 'Direction' may refer to the thematic or disciplinary orientation of the conference.

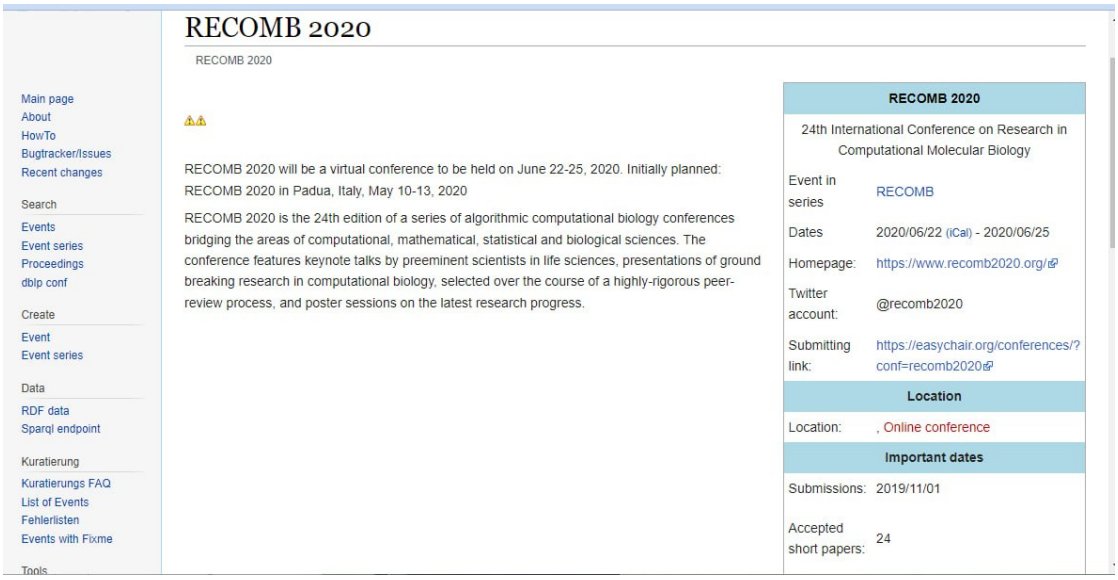


Figure 3: General view of OpenResearch conference platform.

Table 1
A general description of informational-oriented conferences.

Platform	Aim	Short description
ConfIDent (https://www.confident-conference.org/index.php/Main_Page)	To make scientific events persistently accessible in high quality.	A platform for persistent access to high-quality scientific events.
Conference Index (https://conferenceindex.org/)	To help conference organizers reach potential participants easily.	An indexing platform for posting and finding conferences including details like organization and ticket price table.
OpenResearch (https://www.openresearch.org/wiki/Main_Page)	Making descriptive metadata on conferences and other formats of scientific events permanently accessible through automated processes and scientific curating.	A service for researchers to search for and publish information on scientific events.
WikiCFP (http://www.wikicfp.com/cfp/)	To provide a platform for Calls For Papers (CFP) for international conferences, workshops, meetings, seminars, events, journals, and book chapters.	A Wiki website for Calls For Papers in various fields including computer science, engineering, and more.

Fields like 'Region', 'Reference to registration', 'Mode (form of conduct)', 'Hashtags', 'Address', and 'Academic sphere' with a using in two systems are used less frequently but suggest a nuanced

Figure 4: General view of WikiCFP conference platform.

approach to categorizing and promoting conferences. 'Mode (form of conduct)' could indicate whether the conference is virtual or in-person, an increasingly relevant distinction in the post-pandemic landscape.

Finally, 'Website', 'URL program', 'Status', 'Reference to external id', 'ID', 'Doi', 'Contact' is used least frequently. These may represent specific details that are occasionally sought after or are relevant only in certain contexts, such as scholarly communication or detailed inquiries.

The usage frequency of these fields can serve as an indicator of the priorities and behaviors of conference stakeholders. High-frequency fields are likely to be deemed essential for the discovery, selection, and participation in academic conferences, while lower-frequency fields may represent specialized interests or administrative concerns. This data informs platform developers and conference organizers about which features to emphasize or streamline for better user engagement and operational efficiency.

The fields 'Surname', 'Name', and 'Web Page' are used in 5 systems, emerge as the most commonly used. This suggests a strong emphasis on personal identification and online presence,

Figure 5: General view of Morressier platform.

Table 2
General description of process-oriented conferences.

Platform	Aim	Short description	Number of events	Number of visitors
Morressier	To support the entire pre-publishing journey of scholarly communications.	Supports hybrid and virtual conferences, journal submissions, peer-review workflows, and AI-powered checks.	More than 200 professional and scientific organizations use Morressier	Not stated
EasyChair	To organize research paper submission and review for the scientific community.	A conference management system provides various services including a virtual conference solution.	111106 conferences (as of the page's last update)	4094277 users (as of the page's last update)

which are typically crucial for networking, identification, and accessibility purposes.

In contrast, fields with usage in only one system reflect a diverse array of individual and organizational attributes. 'Type of organization' could be indicative of the structure or sector (e.g., non-profit, corporate, academic) an individual is associated with. 'Profile type (open/closed)' might relate to the privacy settings of a user's profile within the organization's database or

Figure 6: General view of EasyChair platform.

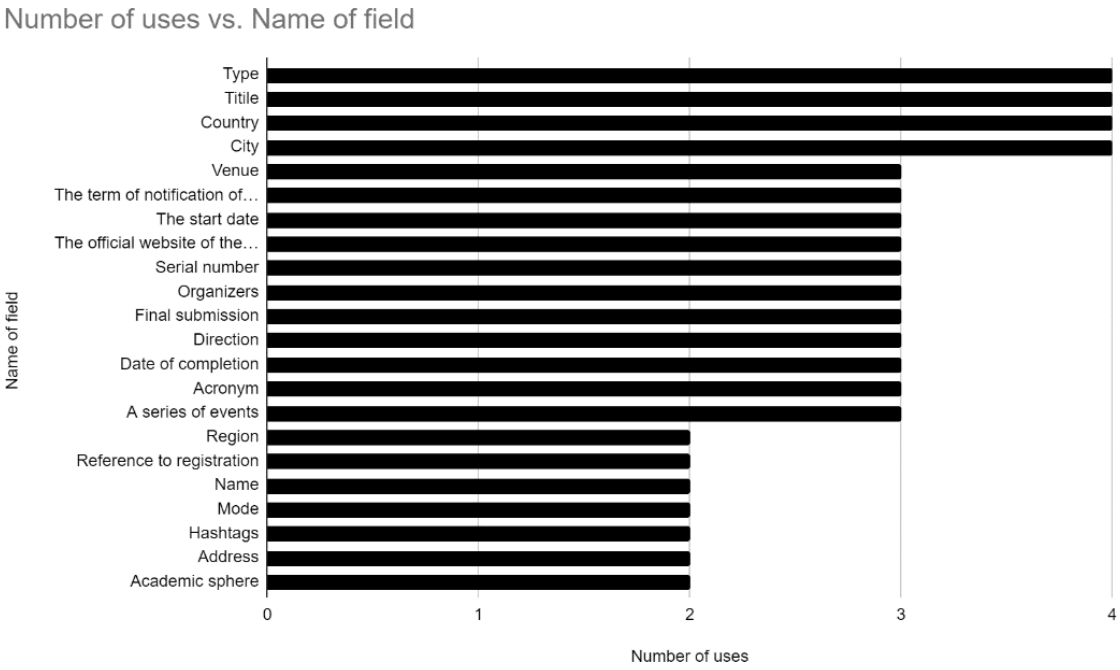


Figure 7: Number of uses of fields of basic metadata of the conferences in conference platforms.

network.

'Position until', 'Position since', and 'Position' fields are indicative of employment status and history. These would be key in understanding tenure, career progression, and current roles within the organization. 'Phone number' and 'Login' are used for communication or system entry purposes.

'Degree', 'Name of the organization', and 'Name of position' provide specific professional and educational details that are important for delineating qualifications and organizational hierarchy. 'Annual income of the organization' is a more specialized field that could be relevant for financial analysis, funding, or economic status reporting within the organization. Considering that annual income of the organization is used for this Morressier system, that is paid system, it seems that this parameter is using to analyze advertising strategies with specific organizations.

The low frequency of these latter fields could suggest that they are either less frequently accessed due to their specificity, or they are utilized in more targeted queries where detailed individual or organizational information is necessary. This distribution indicates a focus on essential identification and communication data over more detailed personal or organizational information in common use cases. Such an analysis can guide the design of information systems, ensuring that frequently used fields are more accessible, while less frequently used data can be structured in a way that is unobtrusive yet available when needed.

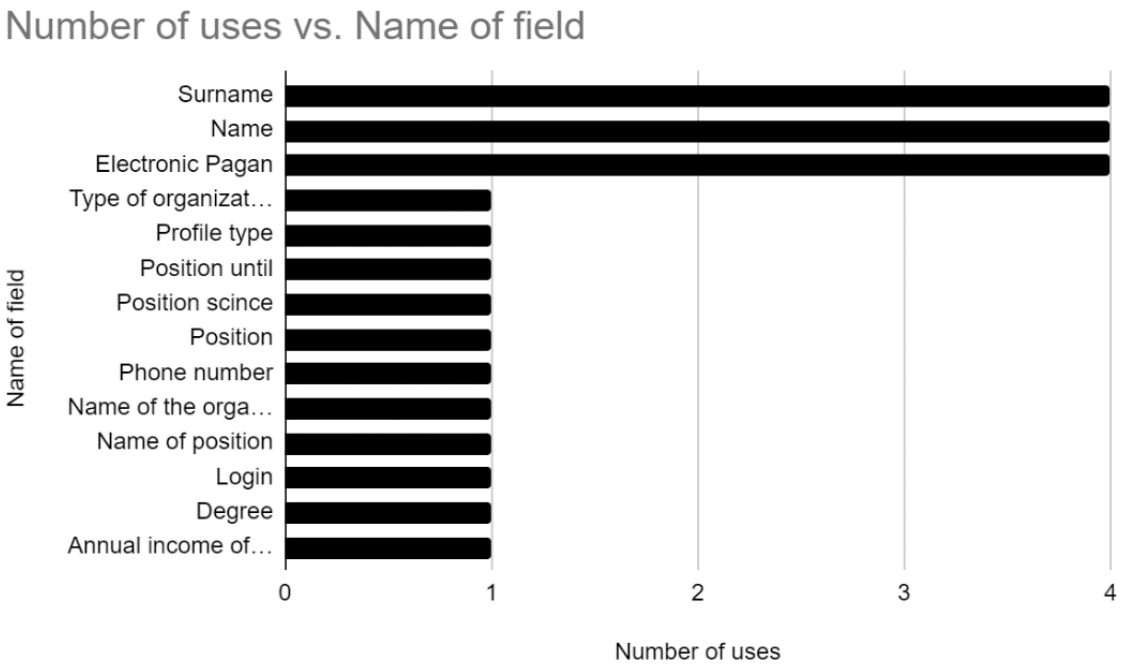


Figure 8: Number of uses of fields of basic metadata of the conferences in conference platforms.

Conference management platforms serve as pivotal tools in the orchestration of scholarly meetings, encompassing a wide array of functions tailored to enhance the academic conference lifecycle as well as only for providing information about specific events. An analysis of common features across such platforms reveals a hierarchy of functionalities, often reflected by their

implementation frequency or utility prominence.

The most prevalent function that is implemented in the 4 systems is the search capability. It allows for an intricate querying of conferences based on a multitude of criteria including geographic location (country and city), academic discipline, and specific conference titles. This omnipresent feature underscores the critical importance of discoverability and accessibility in academic conferencing, suggesting that users place significant value on the ease with which they can locate relevant events.

Peer reviewing and conference registration features implemented only in two systems that are process-oriented systems, indicative of their essential but secondary status in comparison to the search functionality. Peer reviewing is integral to the academic rigor of conference proceedings, enabling the evaluation and selection of scholarly works for presentation. Concurrently, the registration function encapsulates attendee management and may include financial transactions, emphasizing the operational backbone of the conference experience.

Similarly valued, the submission function is an essential one in the academic exchange, facilitating the contribution of research findings to the conference corpus. This is often paralleled by the feature for viewing the list of events, which is equally important for attendees to navigate the conference program effectively.

Publication of materials is in use in two systems, reflecting the necessity for a platform to act as a repository for conference outputs such as proceedings, abstracts, and papers. This feature is crucial for the dissemination of knowledge post-conference and for ensuring the academic contributions are recorded and made accessible.

Less prevalent, yet still integral to the suite of services offered by these platforms, is the capability to manage a series of events, as well as tools to aid in indexing conferences within recognized scientometric databases like Scopus. These functionalities are used only in one system and are suggestive of a more specialized but vital role in enhancing the conference's reach and academic credibility.

This tiered analysis of platform functionalities illustrates a direct correlation between the numerical values and the perceived utility of each feature. The popularity of these functions suggests a prioritization aligned with the primary needs of conference attendees and organizers, which includes finding and participating in relevant academic gatherings, ensuring the quality of content, and the broad dissemination of scholarly work.

'Internal ID' is used in all systems, suggests it is the primary means of identifying records within the system. This high usage underscores the necessity for a unique identifier within an organization's database, enabling efficient management and retrieval of records.

The 'Doi' (Digital Object Identifier) is used in two systems, and it is a widely recognized identifier for electronic documents, which indicates its significant role in the persistent referencing of research articles, datasets, and other academic materials.

'Wikicfp' used in two systems, pointing to their roles as important but less central compared to 'Internal ID' and 'Doi'. 'Wikicfp' might refer to a specific database or index for calls for papers in academic conferences.

The 'Wikidata id', 'ISBN', and 'Dblp' each used only in one system, denoting their specialized use cases. 'Wikidata id' could be a unique identifier for conference series or events within the organization. 'ISBN' (International Standard Book Number) is a familiar identifier for books and other standalone publications, indicating its relevance in academic material management.

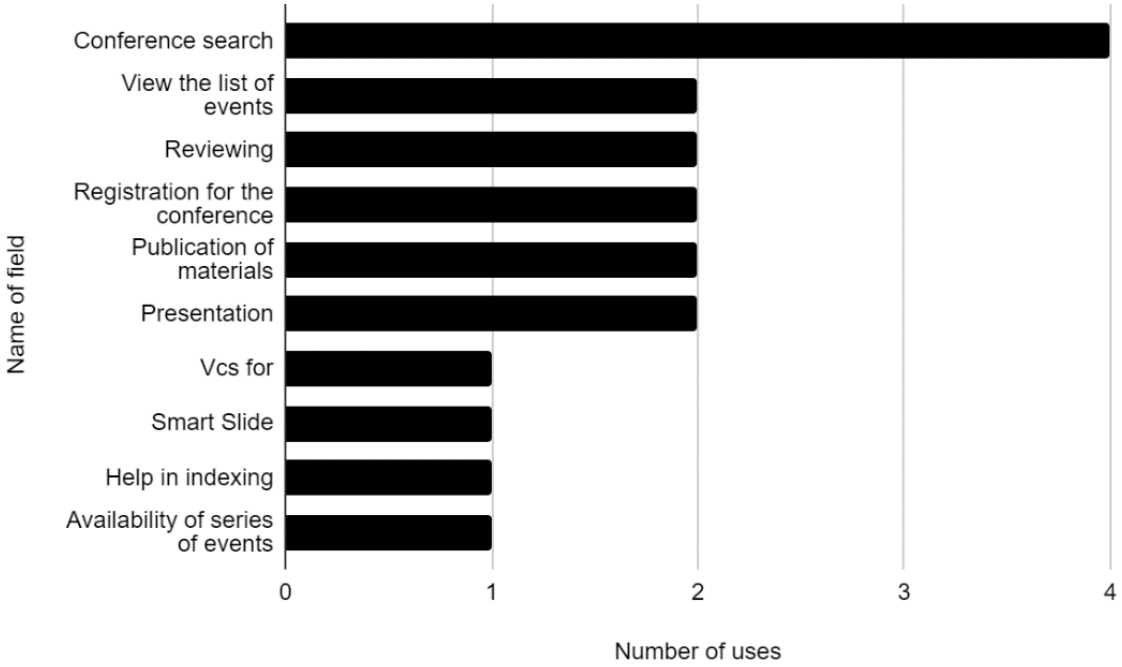


Figure 9: Number of uses of functions in conference platforms.

'Dblp' likely refers to the computer science bibliography, indicating a specific field-related indexing service that supports the citation and tracking of conference proceedings and journals in computing.

This data suggests a hierarchy where internal management and referencing systems (Internal ID), globally recognized identifiers (Doi), and subject-specific databases (Wikicfp, Dblp) play distinct roles in the cataloging, retrieval, and citation of academic work. The lesser frequency of 'ISSN' and 'Wikidata ID' may point to a more contextual application, while 'Doi' and 'ISBN' have broader applications across disciplines and publication types.

3.3. Ontologies as a tool to provide flexible data structures for conferences

Ontologies can ensure flexible data structure that gives an opportunity to both, modify it for each separate conference and modify approach in general when it will be required. CIT Polyhedron is one of the most promising ontology-generating systems that provides generation ontologies and visualization of graphs. Modern ontology systems can make an effective decision-making [14, 15]. It can be effectively used in the in the field of science [16, 17, 18]. It could provide structured view of graph 11 and processing view of table 11. The main advantages of it are the high speed of providing such a system, the possibility to use exchange fields to fill the ontology and interoperability between different knowledge fields.

An ontology-based approach for organizing conference data could counteract the proliferation of predatory and low-quality conferences. This is due to it will promote healthy competition among conference providers by mandating the completion of structured fields. Similar func-

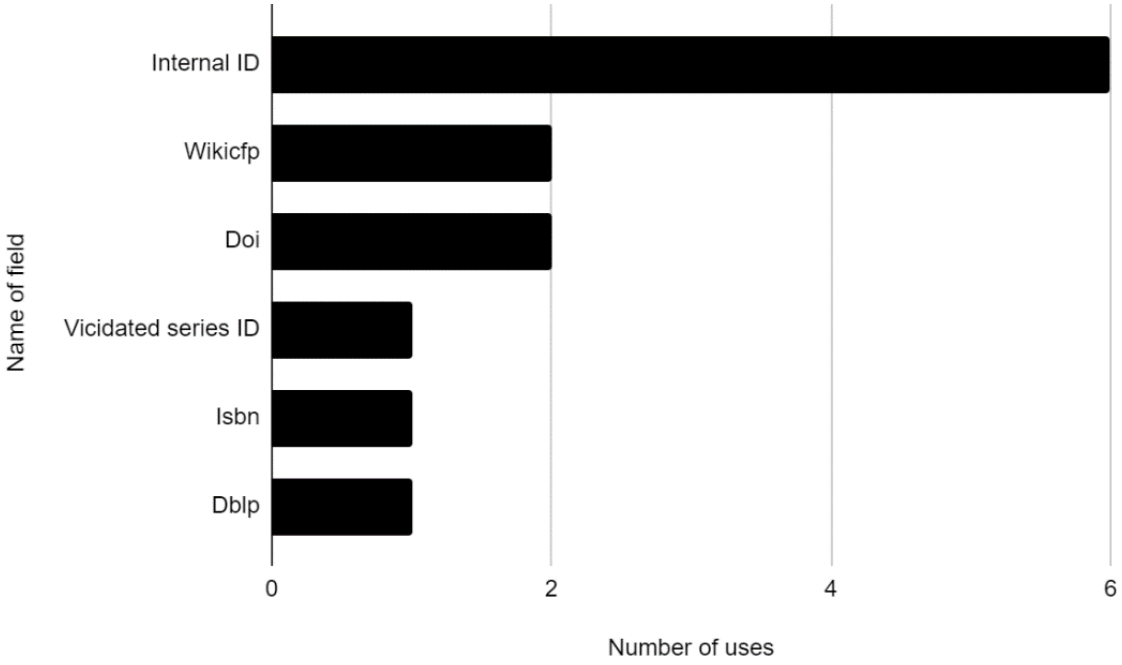


Figure 10: Number of uses of identifiers in conference platforms.

tionalities might be integrated into various systems, but ontology stands out due to its superior level of structuring. An ontology-based system leverages the full potential of consolidating informational resources within the CIT Polyhedron framework. This integration not only enhances data accessibility and organization but also contributes to the overall robustness and efficiency of the system.

4. Conclusion

Conference management platforms are indispensable in the effective management and facilitation of academic conferences. These platforms embody a suite of features that cater to various stages of the conference lifecycle, from initial discovery to post-conference knowledge dissemination. An in-depth analysis of these functionalities reveals a tiered system of importance. At the forefront is the sophisticated search capability, deemed most critical for its role in enabling users to find relevant conferences easily. Subsequent features include peer reviewing and registration, which are crucial for maintaining academic rigor and managing conference operations, respectively. Additionally, the submission of research and the publication of conference materials are fundamental for the scholarly exchange of ideas. Emerging technologies such as VCS for presentation version control and Smart Slide for enhancing presentations are gaining importance. At the base of the hierarchy are internal management and referencing systems, which ensure efficient record management and retrieval. The use of globally recognized identifiers like DOI, along with subject-specific databases, supports the accurate cataloging and citation of academic work. The collective use of these features aims to not only meet the

Figure 11: The use of ontologies in form of graph to systemize conferences.

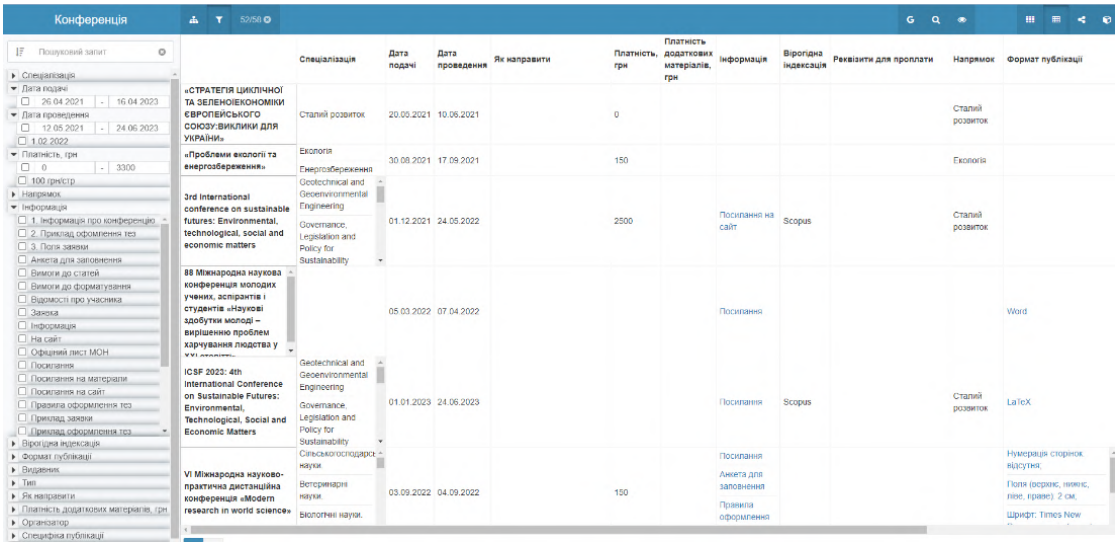


Figure 12: Representing of metadata of the conferences.

primary needs of conference attendees and organizers but also to extend the conference's reach and uphold its academic standing

References

[1] J. McCrostie, Predatory Conferences: A Case of Academic Cannibalism, International Higher Education (2018) 6–8. doi:10.6017/ihe.0.93.10425.

[2] M. Chartier, The Alarming Rise of Predatory Conferences, Eos 103 (2022). doi:10.1029/2022EO220449.

[3] D. Pecorari, Predatory Conferences: What Are the Signs?, Journal of Academic Ethics 19 (2021) 343–361. doi:10.1007/s10805-021-09406-4.

[4] S. Ibrahim, A. Saw, The Perils of Predatory Journals and Conferences, Malaysian Orthopaedic Journal 14 (2020) 1–6. doi:10.5704/MOJ.2007.003.

[5] Z. Stevic, I. Dalic, D. Pamucar, Z. Nunic, S. Veskovic, M. Vasiljevic, I. Tanackov, A new hybrid model for quality assessment of scientific conferences based on Rough BWM and SERVQUAL, Scientometrics 119 (2019) 1–30. doi:10.1007/s11192-019-03032-z.

[6] W. S. Martins, M. A. Gonçalves, A. H. F. Laender, G. L. Pappa, Learning to assess the quality of scientific conferences: a case study in computer science, 2009, p. 193–202. doi:https://doi.org/10.1145/1555400.1555431.

[7] P. Laplante, J. Rockne, P. Montuschi, T. Baldwin, M. Hinchey, L. Shafer, J. Voas, Quality in conference publishing, IEEE Transactions on Professional Communication 52 (2009) 183–196.

[8] G. Berndtsson, M. Folkesson, V. Kulyk, Subjective quality assessment of video conferences and telemeetings, IEEE, 2012, pp. 25–30. doi:10.1109/PV.2012.6229740.

[9] S. Hagemann-Wilholt, C. Hauschke, M. Plank, Confident – An open platform for FAIR conference metadata, Grey Journal 16 (2020) 95–100.
[10] E. Bentlage, L. Roekens, L. Bußmann, M. Brach, Planning and Conducting an Online Conference at the time of COVID-19 : Lessons Learned from EGREPA 2021, BioRxiv (2021) 1–11. doi:10.1101/2021.09.19.460976.
[11] A. Iana, S. Jung, P. Naeser, A. Birukou, S. Hertling, H. Paulheim, Building a Conference Recommender System Based on SciGraph and WikiCFP, in: M. Acosta, P. Cudré-Mauroux, M. Maleshkova, T. Pellegrini, H. Sack, Y. Sure-Vetter (Eds.), Semantic Systems. The Power of AI and Knowledge Graphs, Springer International Publishing, Cham, 2019, pp. 117–123. doi:10.1007/978-3-030-33220-4_9.
[12] S. Zharinov, The role of the library in the digital economy, Information Technology and Libraries 39 (2020). doi:10.6017/ITAL.V39I4.12457.
[13] Y. B. Shapovalov, V. B. Shapovalov, A. G. Zharinova, S. S. Zharinov, I. O. Tsybenko, O. S. Krasovskiy, An academic events sub-system of the URIS and its ontology representation to improve scientific usability and motivation of scientists in terms of European integration, CEUR Workshop Proceedings 3374 (2023) 130–140. URL: https://ceur-ws.org/Vol-3374/paper10.pdf.
[14] O. Y. Stryzhak, V. Horborukov, V. Prychodniuk, O. Franchuk, R. Chepkov, Decision-making System Based on The Ontology of The Choice Problem, Journal of Physics: Conference Series 1828 (2021) 012007–0. doi:10.1088/1742-6596/1828/1/012007.
[15] O. Stryzhak, V. Prykhodniuk, M. Popova, M. Nadutenko, S. Haiko, R. Chepkov, Development of an Oceanographic Databank Based on Ontological Interactive Documents, in: K. Arai (Ed.), Intelligent Computing, Springer International Publishing, Cham, 2021, pp. 97–114. doi:10.1007/978-3-030-80126-7_8.
[16] R. A. Tarasenko, S. A. Usenko, Y. B. Shapovalov, V. B. Shapovalov, A. Paschke, I. M. Savchenko, Ontology-based Learning Environment Model of Scientific Studies, CEUR Workshop Proceedings 3083 (2022) 43–58. URL: https://ceur-ws.org/Vol-3083/paper278.pdf.
[17] Y. B. Shapovalov, V. B. Shapovalov, A Taxonomic Representation of Scientific Studies, CEUR Workshop Proceedings 3013 (2021) 353–360. URL: https://ceur-ws.org/Vol-3013/20210353.pdf.
[18] L. Globa, R. Novogrudskaya, B. Zadoienko, O. Y. Stryzhak, Ontological Model for Scientific Institutions Information Representation, IEEE, 2020, pp. 255–258. doi:10.1109/PICST51311.2020.9467984.

Dynamic system analysis using telemetry

Oleh V. Talaver[1], Tetiana A. Vakaliuk[1,2,3,4]

[1]Zhytomyr Polytechnic State University, 103 Chudnivsyka Str., Zhytomyr, 10005, Ukraine
[2]Institute for Digitalisation of Education of the NAES of Ukraine, 9 M. Berlynskoho Str., Kyiv, 04060, Ukraine
[3]Kryvyi Rih State Pedagogical University, 54 Gagarin Ave., Kryvyi Rih, 50086, Ukraine
[4]Academy of Cognitive and Natural Sciences, 54 Gagarin Ave., Kryvyi Rih, 50086, Ukraine

Abstract

In the modern world of software development, the topic of distributed solutions implementation has become quite common, due to the flexibility it brings to big companies. The downside is that when developing such systems, especially when it comes to many teams, global design problems may not be obvious and lead to a slowdown in the development process or even problems with the location of errors or degradation of overall system performance. In addition, the timely reaction to system degradation is complicated by the distributed nature of the architecture, while manually configuring rules for reporting problematic situations can be time-consuming and still incomplete, automatic detection of possible system anomalies will give engineers (especially Software Reliability Engineers) the focus on problems. For this reason, applications that can dynamically analyze the system for the problems have great potential. Currently, the topic of using telemetry for system analysis is actively studied and gaining traction, so further research is valuable. The aim of the work is to theoretically and practically prove the possibility of using telemetry for the analysis of a distributed information system, detection of harmful architectural practices and anomalous events. To do this, firstly, a detailed overview of the problems related to the topic and the feasibility of using telemetry is provided, the next section briefly describes the history of the monitoring systems development and the key points of the latest OpenTelemetry standard. The main part includes an explanation of the approach used to collect and process telemetry, a reasoning behind the usage of Neo4j as a data storage solution, a practical overview of graph theory algorithms that help in the analysis of the collected data, and a description outlining how the PCA algorithm is employed to detect unusual situations in the whole system instead of individual metrics. The results provide an example of using the software presented in combination with Neo4j Bloom to visualize and analyze the data collected over a period of several hours from the OpenTelemetry Demo test system. The last section contains additional remarks on the results of the study.

Keywords

distributed systems, microservices, dynamic analysis, architectural smells, anti-pattern, visualization, telemetry, anomalies, Open Telemetry, graph theory, statistical analysis

1. Introduction

In recent years, distributed architectures such as microservices have received a lot of attention and popularity due to the opportunities that the architectural pattern opens up in terms

CS&SE@SW 2023: 6th Workshop for Young Scientists in Computer Science & Software Engineering, February 2, 2024, Kryvyi Rih, Ukraine
olegtalaver@gmail.com (O. V. Talaver); tetianavakaliuk@gmail.com (T. A. Vakaliuk)
http://acnsci.org/vakaliuk/ (T. A. Vakaliuk)
0000-0002-6752-2175 (O. V. Talaver); 0000-0001-6825-4697 (T. A. Vakaliuk)
© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

CEUR-WS.org/Vol-3662/paper01.pdf
CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

of optimization, technology stack diversification, and more [1]. Distributed systems, when built correctly, simplify the development process when many teams are involved, reduce the complexity of changes or the dependence of teams on each other, and speed up development. The additional complexity that microservices bring [2] slows down the development of such a system, while the general advantage of technological heterogeneity only increases the effort required to maintain the codebase [3, 4, 5]. Therefore, there is a requirement for more tools to analyze the system and respond to problems. The development of distributed information systems requires more effort, especially when it comes to monitoring the entire system, finding problematic areas [2], because, unlike a monolith, such a system has many components developed in parallel, which may have structural flaws [6, 7], also referred to as architectural smells – design decisions that hinder maintainability and extensibility. For global problems location, an analysis of the information system is often conducted to find and quickly address such shortcomings [8]. A couple approaches exist, such as static analysis of the codebase of each of the system components or analysis of the system logs. Both options are quite complex, because they require adjustment for each individual system, technology, programming language, but the static approach, unlike the dynamic one, allows you to analyze the system without the need to run the whole system, which allows you to correct some local code smells, but not the problems of the system as a whole, due to the low accuracy and insufficient information about the runtime behavior [9]. At the same time, dynamic analysis is based on the information gathered in runtime, therefore presents more accurate representation of the system utilization. The most prominent approach of dynamic analysis is through usage of telemetry that combines three pillars of system observability: logs, metrics, and traces. Therefore, the purpose of the work is the theoretical and practical substantiation of the possibility of using OpenTelemetry standard for the purpose of analyzing a distributed information system: detecting and quickly responding to harmful architectural practices and anomalous events. Next, we outline the tasks:

- research of state-of-art approaches of telemetry analysis;
- modeling extract, transform and load (ETL) and further telemetry analysis process;
- analysis of the received data to identify harmful practices and anomalies.

2. Theoretical background

The topic of system observability is far from new. In the world of distributed systems, Google is considered a pioneer in the study of the topic of observability. In 2010 Google engineers published paper called “Dapper – a Large-Scale Distributed Systems Tracing Infrastructure” [10] which prompts the emergence of the first systems for request traces visualization: Jaeger and Zipkin. However, these applications solved the same problem while being incompatible causing vendor lock, so over time, the development of the OpenTracing [11] standard begun. The new standard provides a layer between the application and monitoring systems to track and collect requests. This standard didn’t solve the whole problem, so OpenCensus standard was later developed to focus on system metrics and logs collection, but also included an alternative implementation of traces collection, which ultimately created more problems, as developers now had to choose between the two standards. For this reason, in 2019, both standards were combined into OpenTelemetry [12] to solve the following tasks:

- gathering traces, metrics, and logs in one place;
- finding anomalies through charts;
- finding the location and cause of anomalies through the review of problematic request traces.

Also, at that time, quite a lot of applications helping in a system monitoring had already been presented on the market, for this reason, one of the tasks was to maintain compatibility with them, so the latest standard provides a specification that describes approaches for metrics collection, conventional descriptions of processes such as interaction using the HTTP protocol, RPC [13] without forcing vendor lock. As a result, OpenTelemetry is currently the most active project of the Cloud Native Computing Foundation [14].

One of the main notions introduced in the standards is telemetry – a set of metrics, logs and, most importantly, traces [15], which in the case of our topic can be used to build a system model in the form of a directed graph [9] and later used to analyze and identify bad practices and problem areas of the system. The OpenTelemetry standard is relatively new, so there is still active research of the possible use cases, but the main area of use is the visualization of requests (figure 1) with the ability to search for problematic areas, for example, the cause of poor service performance or the root cause of an incorrectly working business process [15].

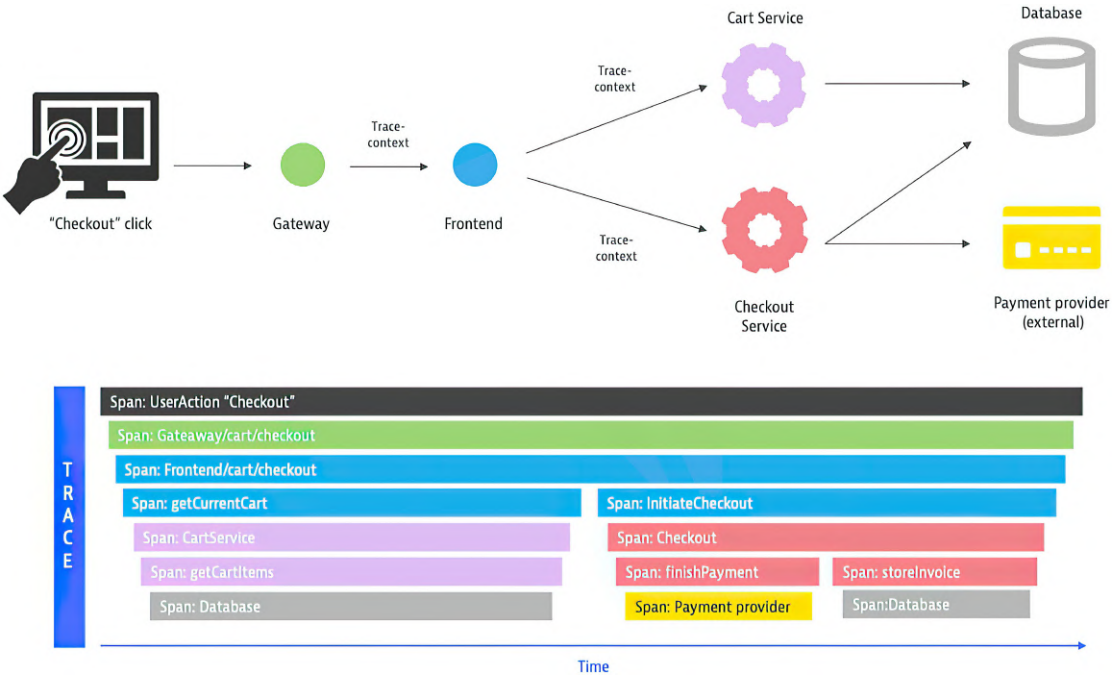


Figure 1: Request trace visualization.

The idea of using telemetry to improve the structure of a system can be traced back to several research papers released in recent years [16, 17, 18], and has a fairly small list of problems that can be identified, which opens up opportunities for further study of this topic [8].

3. Methods

3.1. Defining data and storage for architectural smells detection

The proposed analytical system receives a constant stream of telemetry data, aggregates it by updating the system model in the form of a directed graph stored in a graph database management system (DBMS). After that, the model can be used for analysis, searching for structural anti-patterns.

Constructing the system's graph model involves processing traces of requests (figure 2). Once they are received, the process creates or updates information about available resources (services, storages, proxies) and stores information about changes in the storage. operations available in the service (operation) and individual sub-requests (hop).

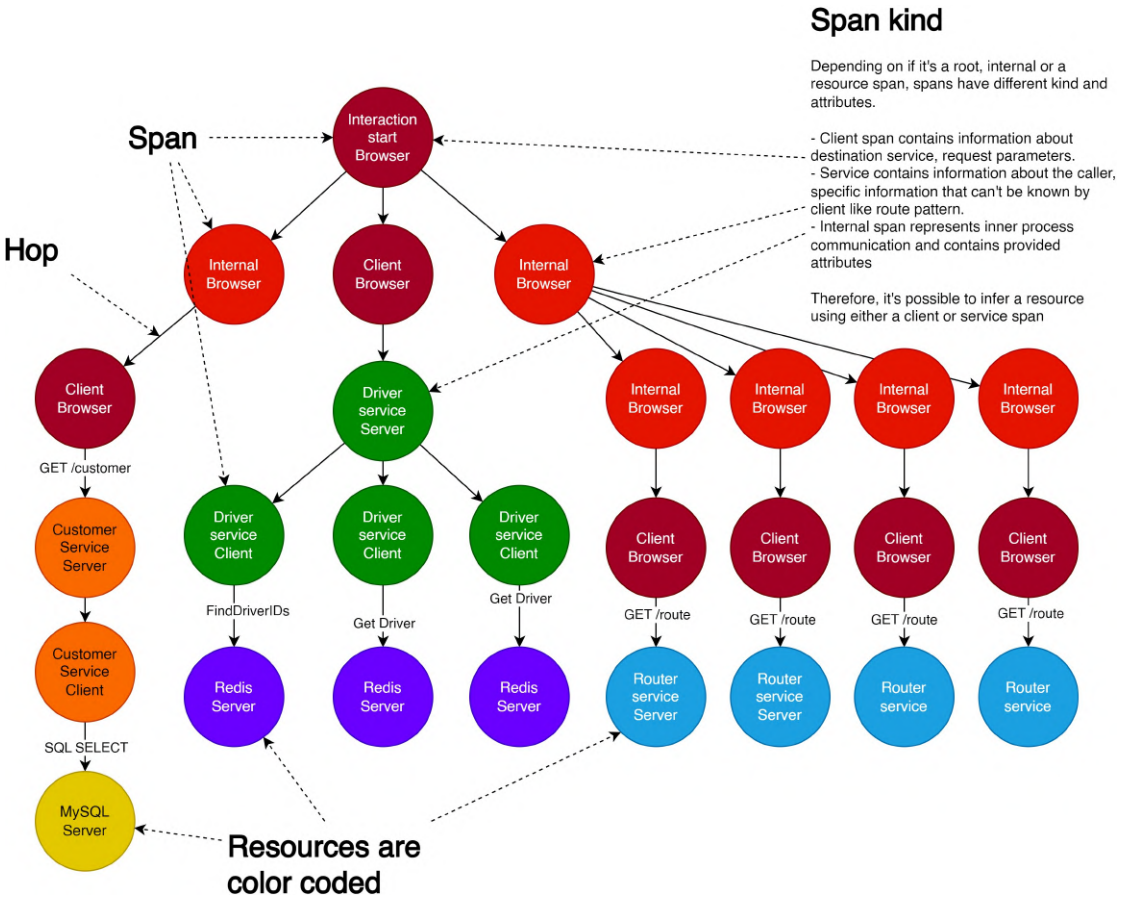


Figure 2: Example of a request tree.

To build a system graph for further analysis, the data storage must have the following information:

- *resources* are interacting components of the system. A resource must have a name, type (service, storage), date of creation and last use;

- *operations* are defined by one resource and called by other ones; have statistics on the number of calls, errors, the last date of creation, use;
- *calls* – connections between a resource and an operation. They has the date of creation, last use, type (synchronous, asynchronous), number of errors.

Neo4j was chosen as the storage of the system model since it physically stores data as a graph, which makes it possible to use graph traversal algorithms to find bad practices in the system, namely:

- clustering coefficient – measures the degree of vertex connectivity; will help show service groups in the system [19];
- degree centrality – measure the number of connections between vertices; makes it possible to calculate the affinity (coupling) metrics of components in the system [20];
- strongly connected components – finds groups where each vertex is accessible from any other; helps to identify cyclic dependencies in the system [21].

The graph DBMS structure is presented in figure 3.

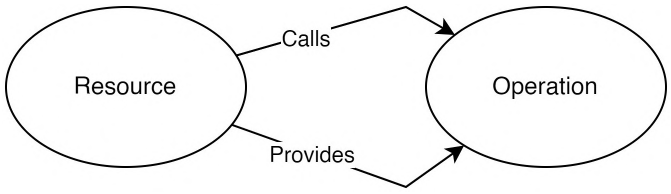


Figure 3: Simplified diagram of the structure of a graph DBMS.

Data storage has two types of nodes: resources and operations. Resources are related to the operations with the "Provides" relation. To show calls, the "Calls" relationship is used, which aggregates statistics for all identical calls from one resource to an operation of another resource.

The ETL process begins with instrumentation of the system – installation of modules for popular libraries that will be collecting the telemetry, manual changes in service code to provide more details of a particular process in the system. Later the telemetry is sent to the OpenTelemetry Collector [22] – a separate modular application developed by the authors of the standard, which allows you to unify the process of collecting, transforming, and exporting telemetry into various popular monitoring systems. There are present numerous available modules, but for this task, it's required to create a custom exporter, which takes batches of traces, extracts necessary data and unloads it into neo4j.

3.2. Methods of anomalies detection

The problem of finding and analyzing anomalies is quite common in computer science and often varies depending on the domain in which the analysis takes place. For example, when reading data from sensors for further analysis, it is important to find and correct outliers. When analyzing a business process, it is sometimes necessary to find unusual events to analyze what

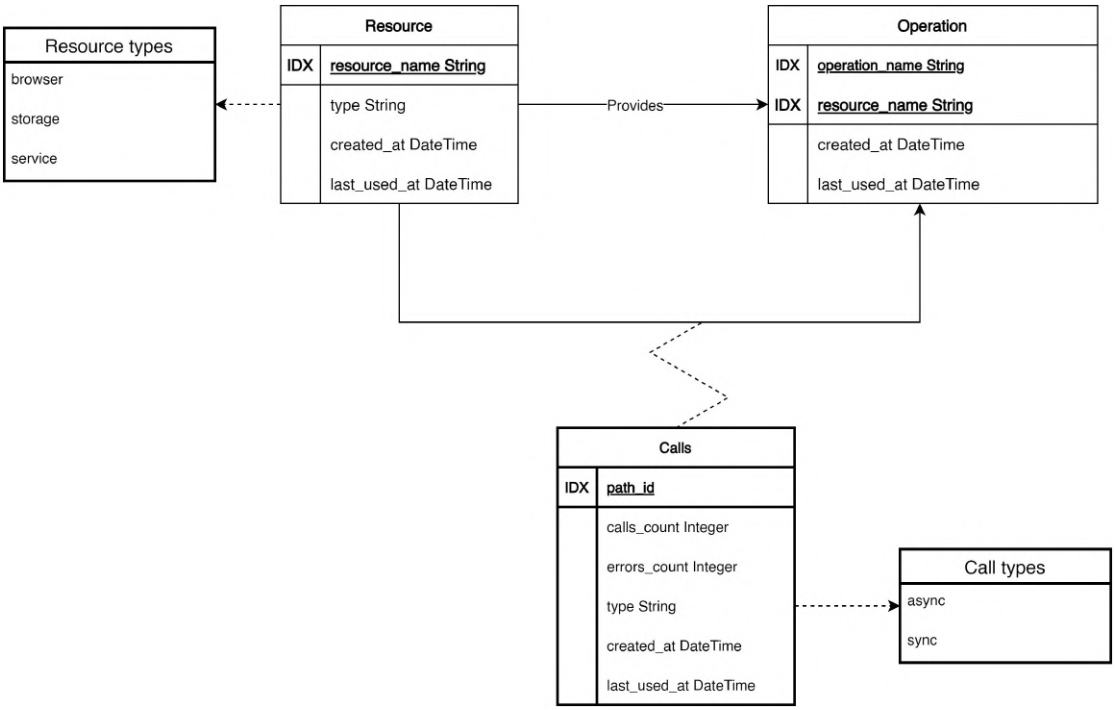


Figure 4: Complete graph DBMS structure diagram.

led to them. In the domain of software reliability engineering, the topic of mean time to detect is one of the most important indicators, because if the problem is found earlier, it is fixed earlier.

Analysis of anomalous changes is already present at least in New Relic, however, it is present at the level of individual services, not the entire system. Although there is not enough data to confirm this, the platform analyzes metrics, including key metrics, using the Exponential Smoothing [23], which is a method of predicting a single variable and, depending on the type, can take into account seasonality [24]. However, it is also possible to find anomalies the opposite way – from a larger scale, using multivariate algorithms, which will be used in this work.

An anomaly is an abnormal situation, essentially defined as a strong difference between expected and actual measurements. Therefore, the process of finding an anomaly includes the process of predicting the value of a certain measurement based on historical data [25].

The problem of finding anomalies in multivariate datasets is quite popular and critical because little to no measurements are univariate [26].

Algorithms are divided into the following training approaches:

- unsupervised – the dataset used for model training does not include labels indicating anomalous situations;
- semi-supervised – the dataset has anomalous situations labeled;
- supervised – the whole dataset is labeled, the least commonly used type of algorithm as it's difficult to get fully labeled data.

Due to the difficulty of obtaining labeled data, unsupervised models are the most popular,

while it's also possible to add the possibility of providing feedback and correction loop when using models for semi-labeled datasets. As part of this work, unsupervised model is reviewed. While "None of the unsupervised methods is statistically better than the others" [25], which is due to the complexity of training on unlabeled data in which extra parameters only interfere, it was decided to choose Principal Component Analysis (PCA) – a statistical method of multivariate analysis used to identify the main structural components in a dataset. The main goal of PCA is to reduce the dimensionality of data while explaining dataset in as much detail as possible, which, due to the simplicity of the approach, is well suited for multivariate datasets and makes it the most common algorithm.

Essentially, PCA converts the initial correlated variables into new linear combinations called principal components. The first principal component is defined in such a way that it explains the largest part of the variance of the data. Each successive principal component is chosen in such a way that it is orthogonal to the previous ones and explains the residual variance as possible.

In the case of identifying anomalies in the system, we are interested in the following information:

- calls – number of incoming, outgoing, and internal calls (synchronous and asynchronous when using a queue or other message brokers) with and without errors;
- duration – time spent processing requests.

To obtain the necessary data in the form of metrics, as well as to group collected data, you need to use a special connector component that transforms traces into call and duration metrics. Thus, the collector receives information about the request via the Open Telemetry Protocol (OTLP) then groups, extracts the necessary metrics to later export it.

Metrics for a certain period are collected for each of the components (resources) of the system. Metrics have different types of values, for example, the number of calls has the sum type that is a counter of certain events for a period and in this case is a monotonous sequence, because the number of calls never decreases.

It is also important to note that the metrics are returned as a delta (the value of aggregationTemporality is 1) and not a cumulative value, because we are interested in the number of calls in a certain period, not the absolute value. Each metric can have multiple points that represent different attribute-defined dimensions (dimensions are customizable), so separate counters have been set up for different request types (span.kind) and statuses (status.code).

Calls duration metric snippet:

```
{
  "name": "duration",
  "unit": "ms",
  "histogram": {
    "dataPoints": [
    {
      "attributes": [
      {
```

```
        "key": "service.name",
        "value": {
          "stringValue": "quoteservice"
        }
      },
      ...
      ],
      "startTimeUnixNano": "1685509043760171242",
      "timeUnixNano": "1685509058790174364",
      "count": "1",
      "sum": 0.006665,
      "bucketCounts": [
      "1",
      "0",
      ...
      ],
      "explicitBounds": [
      0.1,
      1,
      ...
      ],
      "exemplars": [
      {
        "timeUnixNano": "1685509058790174364",
        "asDouble": 0.006665,
        "spanId": "ade03fcb73f18048",
        "traceId": "7f6cf387237813d1f3891b5f21b09be2"
      }
      ]
    },
    ],
    "aggregationTemporality": 1
  }
}
```


If we take the call duration metric, then in this case we have a histogram, which is a certain aggregation of values and their distribution over intervals used for easier visualization.

But in this form, we will not be able to use this data. Firstly, all the metrics for individual services are separated and converted to time series to later be combined based on timestamp.

From the intermediate results, you can clearly see the correlation between the different metrics of the system components (figure 5), which is confirmed by a correlation map (figure 6).

The process of identifying anomalies occurs by splitting the data sample into two periods, the first is used to train the PCA statistical model, the second is used to compare with the predicted values obtained from the model and, estimate the error for all and specific metrics.

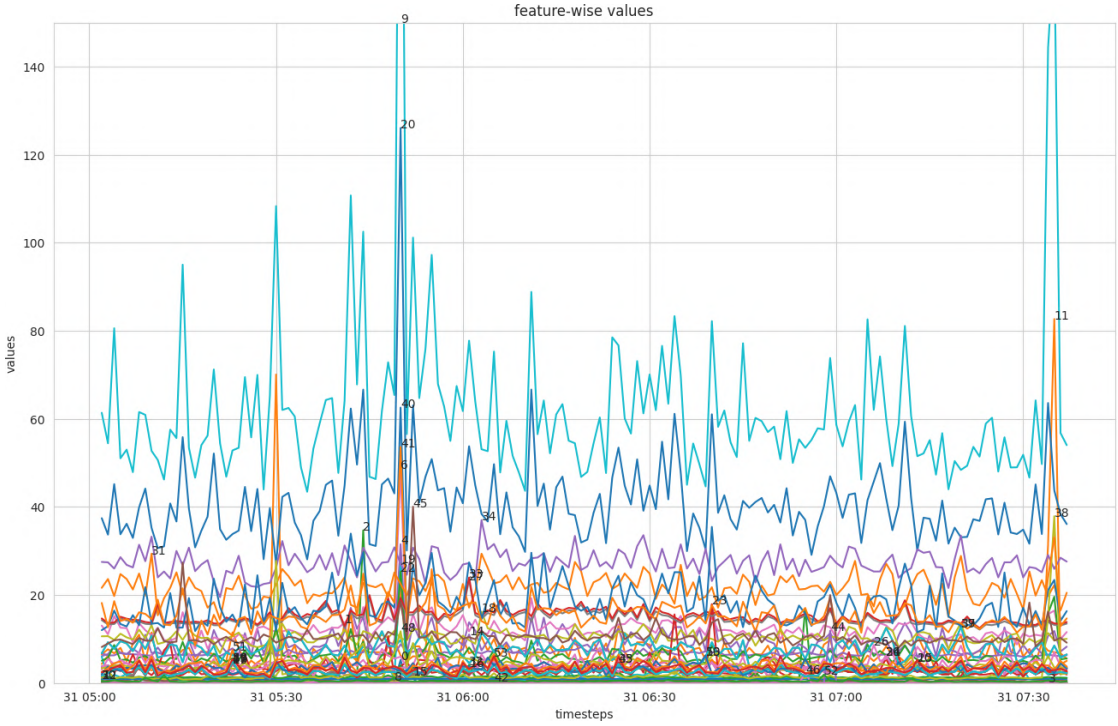


Figure 5: Chart of metric values over time.

4. Results

The OpenTelemetry Demo project was used as a test system [27], which is specially designed for testing applications working with telemetry. This is a distributed system that has components built with different technologies that is automatically loaded using load generator service.

4.1. Visualization of the service graph using Neo4j tools

After some time of running the whole system the graph database has the following data (figure 7). You can see that the graph has many nodes with the type of operation (orange circles), as well as slightly fewer services (purple circles). Between them you can see the “calls” and “provides” relationships depicted as arrows.

To simplify the graph, a function from the APOC library is used [28] for Neo4j in order to visualize the graph projection and show service dependencies (figure 8).

Snippet of a virtual relationship visualization query:

```
MATCH (r1:Resource)-[:Calls]->(:Operation)<-[:Provides]-(r2:Resource)
RETURN r1, r2, apoc.create.vRelationship(r1,'DependsOn',{}, r2) as rel
```

In figure 8 you can see the dependence of the checkout service on many others. To confirm this, let’s use an application called Neo4j Bloom to visualize Local Clustering Coefficient [19] and Degree Centrality [20] algorithms.

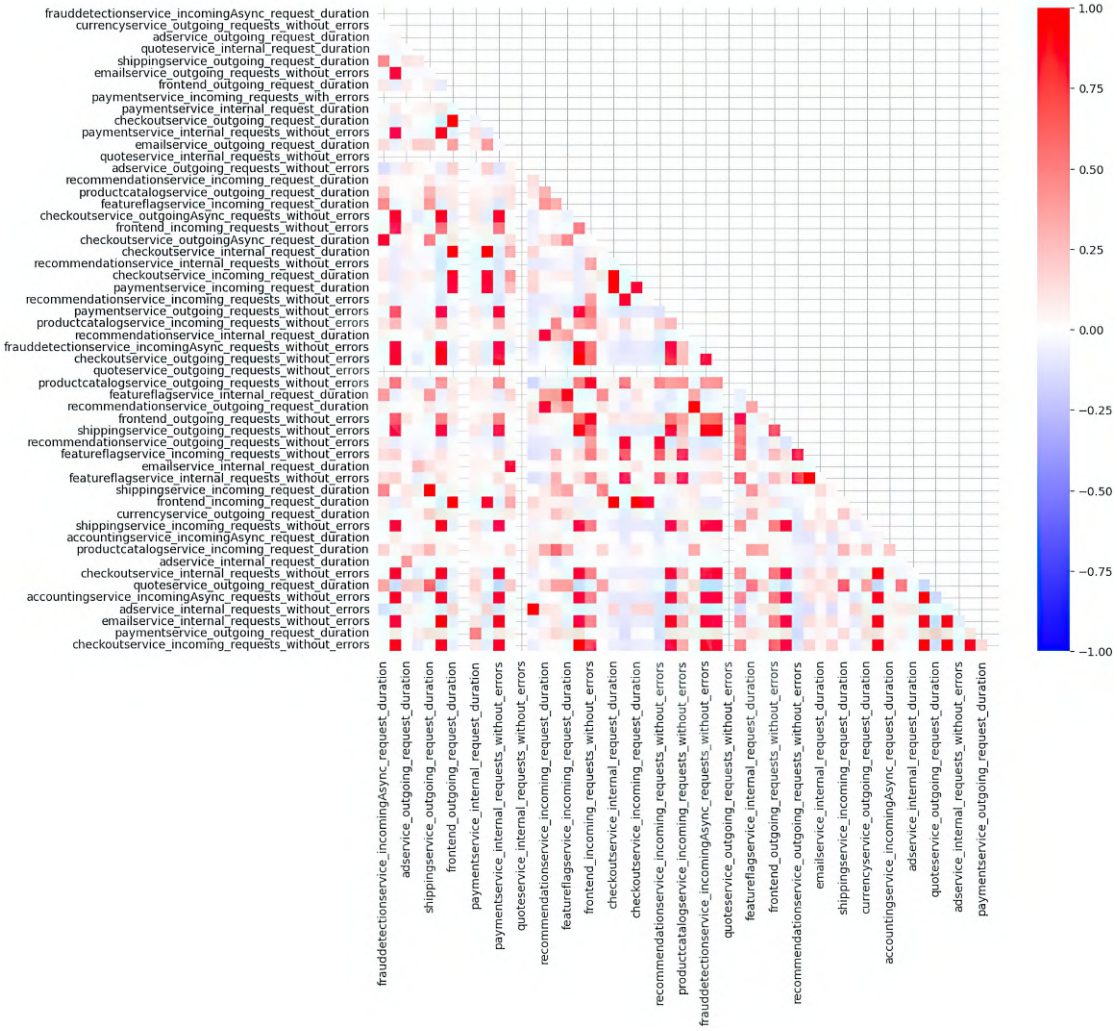


Figure 6: Metrics correlation map.

In the resulting diagram (figure 9) clusters are marked with distinct colors, the dependence of services on peers is indicated with their size. From the diagram it's also clear that the checkout service has many dependencies. This way, you can quickly analyze the architecture of the application, and clearly see parts that must be refactored to prevent a situation where the whole application halts due to a single bottleneck component.

4.2. Time interval anomalies analysis

To detect anomalies, a few hours long time interval was chosen. It has been processed using the PCA algorithm and after receiving errors for each of the time points, a visual analysis can be performed for the presence of spikes in the error values (figure 10).

As you can see in the plot, between 6:50 a.m. and 7 a.m., there were some changes that

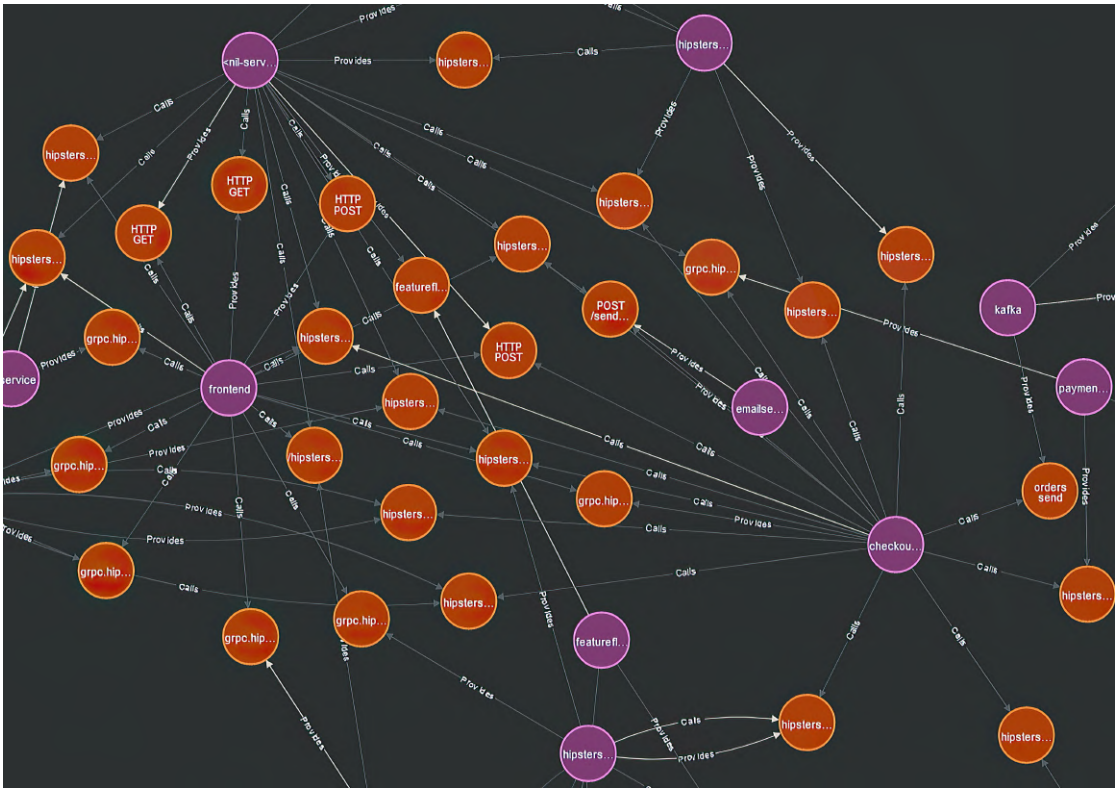


Figure 7: Visualization of the full graph of services, operations and connections between them using Neo4j Browser.

led to a relatively big error. From the error graph for each of the features, it can be seen that feature 44 (accountingservice_incomingAsync_request_duration) is involved in this error, so by conducting a more detailed analysis of the values of this metric, we can see that all values are kept near 0, while there is an outlier with a value of about 12 (figure 11).

5. Discussion

Compared to static analysis approaches, dynamic analysis allows you to see the real picture of the entire system, all possible query paths that are used, accurately indicate the components that cause a problem in the performance of the system at particular moment, in contrast to static analysis of individual modules, which is better suited for the tasks of identifying code smells. Telemetry, in turn, allows you to combine all key indicators and add the additional context that allows you to get more information for analysis.

The practical use of a simple statistical unsupervised PCA algorithm has demonstrated the possibility of using such a model to identify anomalies, which can greatly simplify the work of engineers, because instead of looking on dozens of charts and responding to user messages in support, this statistical analysis suggests the occurrence of anomalous situations in the system automatically. When compared with the approaches of analyzing each metric of the system

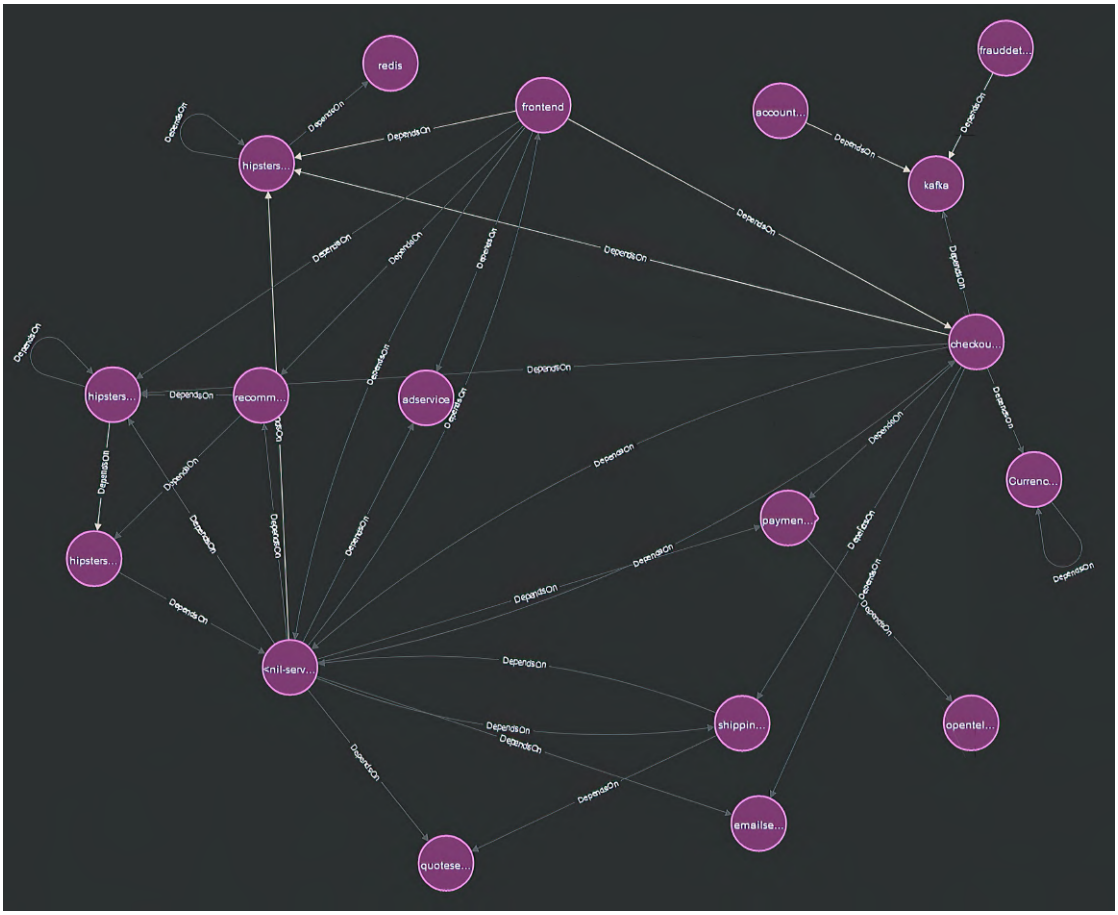


Figure 8: Visualization of the dependency graph in Neo4j Browser.

separately (using appropriate statistical methods, for example, those used in NewRelic [23]), this method gives the general picture, allowing you to understand the situation in the entire system, but also provides the cause of the problem. Compared to supervised algorithms, especially neural networks [29, 30], using the proposed method removes the need to retrain the model to adapt to normal changes (e.g., a natural increase in the number of users of the system), because the analysis takes place in a certain window, although undoubtedly this window should be of a particular size to cover a sufficient amount of data for training and analysis and at the same time not be too sensitive to seasonal changes (for example, activity during the day vs. activity at night), which needs to be tested and determined on a real system.

6. Conclusions

The paper discusses the use of telemetry for dynamic analysis of the system for anomalous events and architectural smells detection.

An analysis of the problems related to distributed systems development was carried out, which

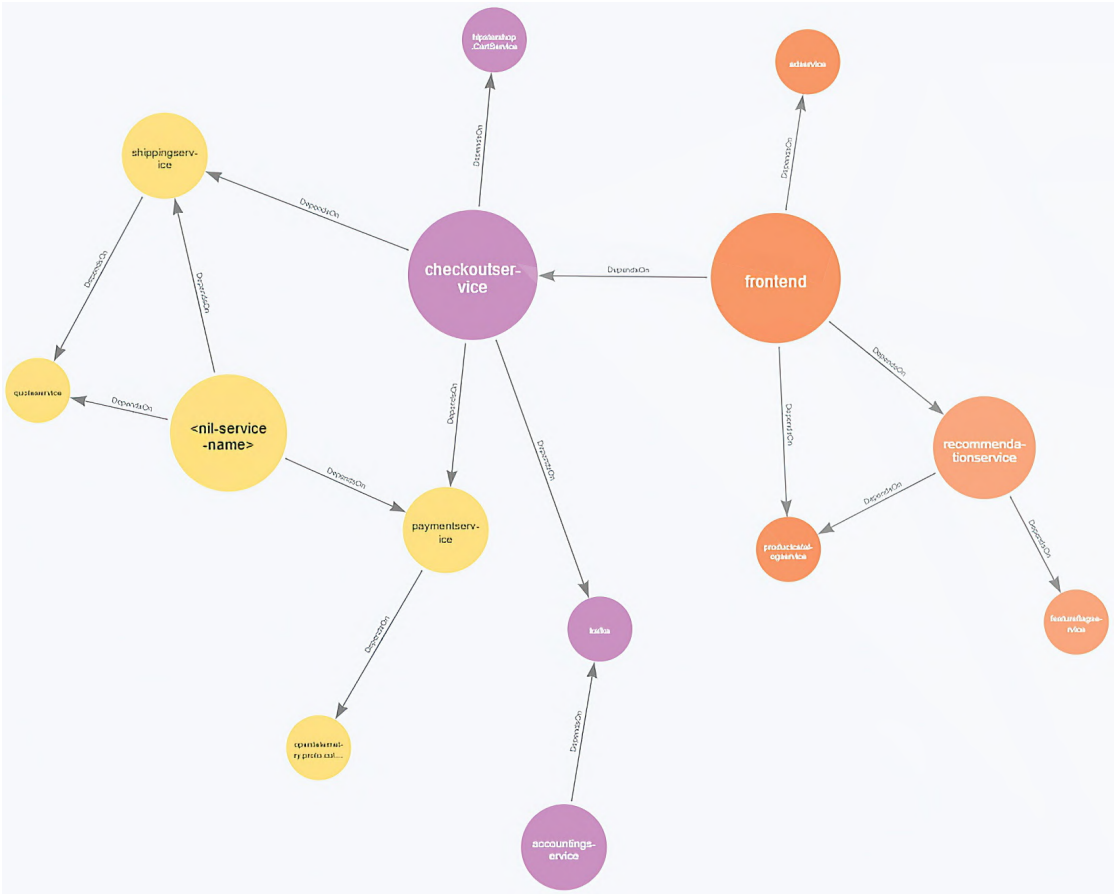


Figure 9: Visualization of the dependency graph of services considering clustering and centrality algorithms in Neo4j Bloom.

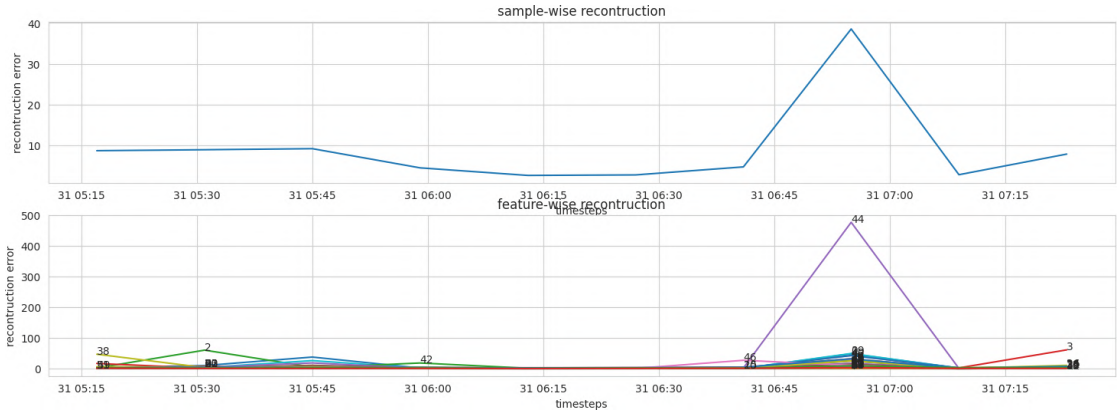


Figure 10: The result of displaying the data reconstruction error for all and individual metrics.

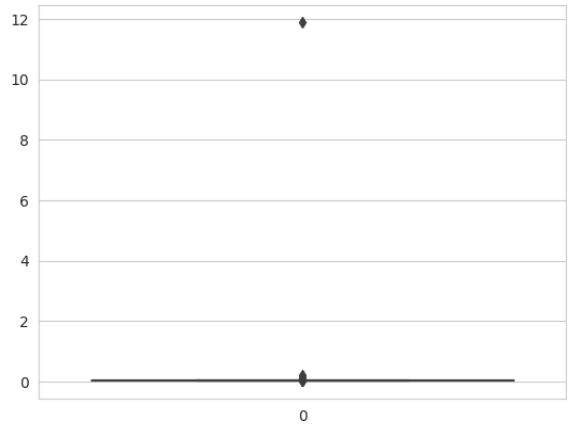


Figure 11: Box plot of values accountingservice_incomingAsync_request_duration.

made it possible to determine the need for applications for monitoring and rapid response to problems in a large system. Studies on the use of telemetry for dynamic system analysis, which have been published in recent years, have shown the potential of this approach. The history of the system monitoring topic development and the main aspects of the latest OpenTelemetry standard were reviewed.

Later, the main data flows that are required for analysis were identified, and a model of a graph DBMS was built. The model includes the following entities: operations, resources and relationships that determine the direction of resource dependence and ownership of operations. After that, the extraction, processing and unloading telemetry using the OpenTelemetry Collector was reviewed. The main types of anomaly detection algorithms were studied and the multivariate PCA statistical method was chosen for the analysis of unlabeled telemetry data. A custom component of the collector application was developed to transform and insert information into Neo4j datastore. The necessary features to be used are defined, as well as the process of collecting appropriate the number and the duration of calls within the system to find anomalies. An algorithm for collecting metrics was described. The next part overviewed the method of using a statistical model of PCA to identify anomalies.

In the next part, aggregated graph model was used to analyze architectural smells. Several possible visualizations of the dependency graph using Neo4j, Neo4j Bloom were provided, clustering and centrality algorithms were used to visually identify problem areas in the system architecture. The results of the statistical model based on the Principal Component Analysis algorithm were also analyzed. The accuracy of this model is sufficient to determine anomalous events.

The topic of telemetry usage to find bad architectural practices has potential for further development [18], after all, the collected data is enough to determine more complex patterns: too long synchronous and asynchronous requests, long chains of synchronous requests, the presence of too many different technologies in a small system, a large time difference between when an event is published and processed. To provide even more opportunities for analysis, it's possible to enrich system resources with additional metadata indicating belonging to a certain bounded context (to compare de jure and de facto clusters of contexts, to identify situations

of using the same database in different parts of the application) [1], dates of any changes or releases. Also, this data should be displayed on the main map of the system components. Further development of anomaly analysis includes integration with an application performance monitoring (APM) system, the ability to configure threshold values for reconstruction error, adding custom metric streams for analysis and finally testing on a real system with a comparison with existing approaches.

If we consider other topics of telemetry usage, we cannot omit the topic of analyzing individual use cases in the application, which are sets of requests that go through a bunch of services and have variations depending on some stored state of the application, the analysis includes both the ability to evaluate performance, the number of errors of a particular use case, the ability to subscribe a certain development team to updates for a quick response in case of anomalous situations, and ability to view changes, including performance, between different releases.

References

[1] P. D. Francesco, I. Malavolta, P. Lago, Research on Architecting Microservices: Trends, Focus, and Potential for Industrial Adoption, in: 2017 IEEE International Conference on Software Architecture (ICSA), 2017, pp. 21–30. doi:10.1109/ICSA.2017.24.

[2] M. Söylemez, B. Tekinerdogan, A. Kolukısa Tarhan, Challenges and Solution Directions of Microservice Architectures: A Systematic Literature Review, Applied Sciences 12 (2022) 5507. doi:10.3390/app12115507.

[3] A. Bucchiarone, N. Dragoni, S. Dustdar, S. T. Larsen, M. Mazzara, From Monolithic to Microservices: An Experience Report from the Banking Domain, IEEE Software 35 (2018) 50–55. doi:10.1109/MS.2018.2141026.

[4] M. Villamizar, O. Garcés, H. Castro, M. Verano, L. Salamanca, R. Casallas, S. Gil, Evaluating the monolithic and the microservice architecture pattern to deploy web applications in the cloud, in: 2015 10th Computing Colombian Conference (10CCC), 2015, pp. 583–590. doi:10.1109/ColumbianCC.2015.7333476.

[5] O. V. Talaver, T. A. Vakaliuk, Reliable distributed systems: review of modern approaches, Journal of Edge Computing 2 (2023) 84–101. doi:10.55056/jec.586.

[6] J. Soldani, D. A. Tamburri, W.-J. Van Den Heuvel, The pains and gains of microservices: A Systematic grey literature review, Journal of Systems and Software 146 (2018) 215–232. doi:10.1016/j.jss.2018.09.082.

[7] S. Niedermaier, F. Koetter, A. Freymann, S. Wagner, On Observability and Monitoring of Distributed Systems – An Industry Interview Study, in: S. Yangui, I. Bouassida Rodriguez, K. Drira, Z. Tari (Eds.), Service-Oriented Computing, Springer International Publishing, Cham, 2019, pp. 36–52. doi:10.1007/978-3-030-33702-5_3.

[8] I. Pigazzini, F. A. Fontana, V. Lenarduzzi, D. Taibi, Towards Microservice Smells Detection, in: Proceedings of the 3rd International Conference on Technical Debt, TechDebt '20, Association for Computing Machinery, New York, NY, USA, 2020, p. 92–97. doi:10.1145/3387906.3388625.

[9] T. Cerny, A. S. Abdelfattah, V. Bushong, A. Al Maruf, D. Taibi, Microservice Architecture Reconstruction and Visualization Techniques: A Review, in: 2022 IEEE International

Conference on Service-Oriented System Engineering (SOSE), 2022, pp. 39–48. doi:10.1109/SOSE55356.2022.00011.
[10] B. H. Sigelman, L. A. Barroso, M. Burrows, P. Stephenson, M. Plakal, D. Beaver, S. Jaspan, C. K. Shanbhag, Dapper, a Large-Scale Distributed Systems Tracing Infrastructure, 2010. URL: https://api.semanticscholar.org/CorpusID:14271421.
[11] The OpenTracing Semantic Specification, 2023. URL: https://opentracing.io/specification/.
[12] OpenTelemetry, 2024. URL: https://opentelemetry.io/docs/what-is-opentelemetry/.
[13] OpenTelemetry Semantic Conventions, 2024. URL: https://opentelemetry.io/docs/specs/semconv/.
[14] OpenTelemetry Project Journey Report – 2023, 2023. URL: https://www.cncf.io/reports/opentelemetry-project-journey-report/.
[15] Observability Primer, 2023. URL: https://opentelemetry.io/docs/concepts/observability-primer/.
[16] G. Parker, S. Kim, A. A. Maruf, T. Cerny, K. Frajtak, P. Tisnovsky, D. Taibi, Visualizing Anti-Patterns in Microservices at Runtime: A Systematic Mapping Study, IEEE Access 11 (2023) 4434–4442. doi:10.1109/ACCESS.2023.3236165.
[17] I. U. P. Gamage, I. Perera, Using dependency graph and graph theory concepts to identify anti-patterns in a microservices system: A tool-based approach, in: 2021 Moratuwa Engineering Research Conference (MERCon), 2021, pp. 699–704. doi:10.1109/MERCon52712.2021.9525743.
[18] X. Guo, X. Peng, H. Wang, W. Li, H. Jiang, D. Ding, T. Xie, L. Su, Graph-Based Trace Analysis for Microservice Architecture Understanding and Problem Diagnosis, in: Proceedings of the 28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering, ESEC/FSE 2020, Association for Computing Machinery, New York, NY, USA, 2020, p. 1387–1397. doi:10.1145/3368089.3417066.
[19] Neo4j Local Clustering Coefficient, 2023. URL: https://neo4j.com/docs/graph-data-science/current/algorithms/local-clustering-coefficient/.
[20] Neo4j Degree Centrality, 2023. URL: https://neo4j.com/docs/graph-data-science/current/algorithms/degree-centrality/.
[21] Neo4j Strongly Connected Components, 2023. URL: https://neo4j.com/docs/graph-data-science/current/algorithms/strongly-connected-components/.
[22] OpenTelemetry Collector, 2023. URL: https://opentelemetry.io/docs/collector/.
[23] N. D. Boone, Dynamic Baseline Alerts Now Automatically Find the Best Algorithm for You, 2017. URL: https://newrelic.com/blog/how-to-relic/baseline-alerts-algorithm.
[24] J. Brownlee, A Gentle Introduction to Exponential Smoothing for Time Series Forecasting in Python, 2020. URL: https://machinelearningmastery.com/exponential-smoothing-for-time-series-forecasting-in-python/.
[25] S. Han, X. Hu, H. Huang, M. Jiang, Y. Zhao, ADBench: Anomaly Detection Benchmark, in: S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, A. Oh (Eds.), Advances in Neural Information Processing Systems, volume 35, Curran Associates, Inc., 2022, pp. 32142–32159. URL: https://proceedings.neurips.cc/paper_files/paper/2022/file/cf93972b116ca5268827d575f2cc226b-Paper-Datasets_and_Benchmarks.pdf.
[26] S. Suboh, I. Aziz, S. Shaharudin, S. Ismail, H. Mahdin, A Systematic Review of Anomaly Detection within High Dimensional and Multivariate Data, JOIV : International Journal

on Informatics Visualization 7 (2023) 122. doi:10.30630/joiv.7.1.1297.
[27] OpenTelemetry Demo, 2023. URL: https://github.com/open-telemetry/opentelemetry-demo.
[28] Neo4j APOC Library, 2023. URL: https://neo4j.com/developer/neo4j-apoc/.
[29] S. O. Semerikov, T. A. Vakaliuk, I. S. Mintii, V. A. Hamaniuk, V. N. Soloviev, O. V. Bondarenko, P. P. Nechypurenko, S. V. Shokaliuk, N. V. Moiseienko, V. R. Ruban, Development of the computer vision system based on machine learning for educational purposes, Educational Dimension 5 (2021) 8–60. doi:10.31812/educdim.4717.
[30] I. A. Pilkevych, D. L. Fedorchuk, M. P. Romanchuk, O. M. Naumchak, Approach to the fake news detection using the graph neural networks, Journal of Edge Computing 2 (2023) 24–36. doi:10.55056/jec.592.

Development of a modified genetic method for automatic university scheduling

Ievgen Fedorchenko, Andrii Oliinyk, Tetiana Zaiko, Kyrylo Miedviediev, Yuliia Fedorchenko and Mykola Khokhlov

National University "Zaporizhzhia Polytechnic", 64 Zhukovsky Str., Zaporizhzhia, 69063, Ukraine

Abstract

The paper investigates the problem of optimizing the university class schedule. Sequential and parallel methods for scheduling based on genetic search are developed. The proposed methods use adapted initialization, crossover, and selection operators. The algorithms minimize conflicts and the time interval between classes, taking into account recommendations for time and place. The developed methods contribute to effective planning of the educational process and avoidance of errors in scheduling. A comparative analysis of the classical and modified genetic algorithms is carried out, confirming the faster and more efficient functioning of the modified approach. The modified algorithm is also compared with different operators and parameters of the genetic algorithm to determine the optimal conditions. The obtained results indicate effective methods for improving the quality of the schedule and optimizing the educational process at the university.

Keywords

genetic algorithm, schedule, evolutionary algorithm, classes, classes

1. Introduction

Time and resource management is critical to success in higher education. For students, effective time management is an important factor for successful studies, allowing them to balance studies with other activities such as work, sports or social life. On the other hand, for teachers, an optimal schedule of classes helps increase the productivity and efficiency of their work.

Scheduling is a complex task that requires a lot of resources and planning knowledge. Since there are many constraints, such as following the curriculum and taking into account the needs of students and teachers, automating this process is essential. The scheduling problem can be classified as an NP-complete problem due to the large number of possible solutions, and metaheuristic algorithms are used to solve it. This paper proposes a modified genetic algorithm for efficient university scheduling.

CS&SE@SW 2023: 6th Workshop for Young Scientists in Computer Science & Software Engineering, February 2, 2024, Kryvyi Rih, Ukraine

evg.fedorchenko@gmail.com (I. Fedorchenko); olejnikaa@gmail.com (A. Oliinyk); nika270202@gmail.com (T. Zaiko); kirillmedvedev279@gmail.com (K. Miedviediev); jul.fedorchenko@gmail.com (Y. Fedorchenko); khokhlov@zp.edu.ua (M. Khokhlov)

0000-0003-1605-8066 (I. Fedorchenko); 0000-0002-6740-6078 (A. Oliinyk); 0000-0003-1800-8388 (T. Zaiko); 0000-0003-3200-7306 (K. Miedviediev); 0000-0003-4436-3877 (Y. Fedorchenko); 0000-0001-8272-9847 (M. Khokhlov)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

CEUR-WS.org/Vol-3662/paper05.pdf

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

2. Literature review and problem statement

Implementation of automatic systems for drawing up university timetables is an urgent task, that is of great importance for improving the management of the educational process. Over the last decades, many studies and developments have been carried out in order to optimize the process of the distribution of working time and resources in university educational institutions.

Lukas et al. [1] use a genetic algorithm and a heuristic search to solve the problem of scheduling at a university. The schedule creation methodology is based on genetic algorithms, which are aimed at maximizing the number of successfully planned lesson units in the schedule. This article provides a detailed description of the timetable generation process and takes into account various constraints, such as the availability of teachers, classrooms and days.

Abduljabbar and Abdullah [2] uses the method of genetic algorithms to solve the problem of a complex schedule of classes at a university or college. The authors proposed a system where complex aspects such as class schedules, lecture times, and available classrooms are encoded as binary sequences, and based on these, solutions are generated using genetic operations such as selection, crossover, and mutation.

Alghamdi et al. [3] examines the problem of scheduling classes in universities and explores various methods for its optimization. The article provides an overview of the genetic algorithm optimization method. Genetic algorithms are based on the principles of genetics and natural selection. The basic idea is to select the fittest individuals in a population, which are then recombined or mutated to create new groups. GAs can effectively solve optimization problems that have different parameters or characteristics that cannot be represented mathematically.

The task of creating a schedule is to create an optimal schedule of classes, taking into account various restrictions and requirements. This issue belongs to the class of NP-complete problems because it is a combinatorial optimization problem in which the number of solutions increases with the size of the input data, constraints, and parameters.

The objective function for the scheduling problem defines the main optimization criteria, which include:

- minimizing the number of windows between classes to reduce the downtime of teachers and students;
- uniform distribution of the load during the week to increase work efficiency;
- minimization of conflicts and overlaps of classes, which involves avoiding overlap between classes with joint groups, teachers or classrooms;
- taking into account proposals regarding the schedule to meet the needs of students and teachers.

The objective function calculates the sum of the penalty functions that indicate violations of the constraints when creating the schedule. Each penalty function can have a weight depending on the importance of the corresponding constraint. We use the following notations: n is the number of groups, m is the number of teachers, k is the number of audiences, L is the number of classes, D is the number of school days, U is the maximum number of classes per day, Y is a type of couple, $G = [g_1, g_2, \ldots, g_n]$ is schedule relative to groups, $T = [t_1, t_2, \ldots, t_m]$ is timetable for teachers, $A = [a_1, a_2, \ldots, a_k]$ is the timetable for the classrooms, $S = [s_1, s_2, \ldots, s_l]$ is general schedule, X is the schedule of a separate group, teacher or classrooms.

$f_1(x)$ is a function that calculates windows between classes and has a corresponding weighting factor k_1. The function calculates the difference between the next class and the previous one within one day. Depending on the type of classes, the following differences will be calculated: between two common classes, or between the common and the numerator or denominator, and between two classes in the numerator or denominator. In general, this function can be described as follows (1):

$$f_1(x) = k_1 \sum_{d=1}^{D} \sum_{u=1}^{U} \sum_{y=1}^{Y} x_{d,u,y} - x_{d,u-1,y}, \tag{1}$$

$f_2(x)$ is a function that calculates the overlap of classes and has the corresponding weighting coefficient k_2. This function checks whether classes in a certain group or audience do not overlap. the teacher An overlap is considered when two or more classes are scheduled for the same day of the week and class number, or different classes are held in the same auditorium. It should be noted that if two classes are held in the same time slot, the class type is the numerator and the denominator, respectively, then this is not considered overlapping. Mathematical model of function (2):

$$f_2(x) = k_2 \sum_{i=1}^{L} \sum_{j=i+1}^{L} w_2(x_i, x_j) \tag{2}$$

$w_2(x_i, x_j)$ is a Boolean function that can be described as follows (3):

$$w_2(x_i, x_j) = \begin{cases} 1, & \text{if i=j .} \\ 0, & \text{else.} \end{cases} \tag{3}$$

$f_3(x)$ is a function that calculates compliance with the recommendations regarding the time of the lesson and has a corresponding weighting factor k_3. The objective function will look like this (4):

$$F(S) = \sum_{i=1}^{2} \sum_{j=1}^{n} k_i f_i(g_j) + \sum_{i=1}^{2} \sum_{j=1}^{m} k_i f_i(t_j) + k_2 \sum_{j=1}^{k} f_2(a_j) + k_3 \sum_{j=1}^{l} f_3(s_j) \tag{4}$$

With the help of weighting coefficients, we can determine the importance of compliance with the corresponding restriction. If the value of the coefficient is high, the priority in the algorithm will be to minimize this limitation.

Planning training sessions is a complex combinatorial task, and traditional scheduling methods are not suitable for it. They use heuristic methods, such as genetic algorithms, which, although they do not guarantee the best solution, effectively search for optimal options.

During the analysis of the available methods of solving the problem, it was established that the most suitable algorithm for drawing up the optimal schedule of classes is the genetic algorithm. Therefore, it was decided to develop an adapted and modified genetic algorithm for solving the given problem.

The scheduling algorithm should minimize the value of the defined objective function (4). The work of the algorithm can be represented in the form of the following formula (5):

$$F(S) \to min, \tag{5}$$

where S is the class schedule; $F(S)$ is the objective function.

To evaluate the quality of the results of the algorithm, the objective function described above will be used, which shows the number of violations of the specified restrictions.

3. The purpose and objectives of the research

The *purpose of the research* is to develop a mathematical model for solving the task of creating an optimal schedule of classes for higher educational institutions.

To achieve this purpose, it was necessary to solve the following tasks:

- to develop an adapted classic genetic algorithm for drawing up an optimal class schedule;
- to develop a modified method of the classical genetic algorithm;
- to present an experimental study of the proposed genetic methods.

4. Development of genetic algorithm modification

An evolutionary genetic method was chosen to optimize the objective function. The genetic algorithm gradually approaches the optimal schedule by selecting, combining, and varying possible solutions.

The main steps of the genetic algorithm are [4, 5]:

1) initialization of input data and algorithm parameters;
2) initialization of the initial population;
3) calculation of the adaptability of individuals;
4) crossing of individuals;
5) mutation of individuals in the population;
6) calculation of the suitability of the obtained individuals;
7) selection;
8) continuation of steps until the stop criterion is satisfied.

The initial data for the system was obtained from the semester assignment information from the software department. This document included a list of subjects, a list of groups attending these subjects, the names of lecturers who give lectures or practical classes, and recommendations for the classrooms needed for the classes. The task was to create a schedule for 28 teachers and 40 groups, using 10 classrooms (6 lecture rooms and 4 laboratory rooms).

The input data for the algorithm were:

- a list of classes, each of which included a teacher or teachers, one or more groups, a subject, a type of class, and the number of meetings;
- recommended schedule for classes;

- the number of school days;
- the maximum number of classes per day;
- list of classrooms.

The algorithm has the following parameters;

- population size;
- the number of iterations;
- probability of mutation;
- probability of crossing;
- probability of gene mutation.

Initialization in the genetic algorithm is the process of creating an initial population of individuals that will evolve in the future [6].

Two initialization methods were developed.

The first method is random initialization. Random initialization makes it possible to include various solutions in the initial set, which increases the probability of finding an optimal schedule. For each lesson, the day of the week, session number, type and audience are randomly determined. If the class has a recommended schedule, it will be taken into account [6].

The second method is the initialization of the schedule, modification of the random one: we randomly generate a day of the week, choose an audience, then look for a possible time of the class. If this is the first lesson on the selected day, you can insert it randomly. If there is already a class, then the number is selected in such a way as to minimize violation of the restriction. As a result, we will get a schedule in which there will be no overlap of classes and a reduced number of windows. Using this initialization method will significantly speed up the work of the genetic algorithm.

Crossover is an operation in a genetic algorithm that is used to create a new population. Crossbreeding consists in the exchange of genes between two parental individuals in order to create offspring with a new combination of genes [6].

It was decided to exchange only one random activity in two random schedules. This allows other genes to change in the next generation and preserve beneficial combinations.

The k-point crossover method involves choosing k points on the parental chromosomes, dividing the chromosomes into $k + 1$ segments, and exchanging these segments to create two offspring. An even multiple of k ensures the division of segments into equal parts [6].

Advantages of this method include diversity in offspring, preservation of useful genetic combinations from parents, and prevention of collapse into a local minimum during optimization. However, a large value of k can lead to the loss of useful information from parental chromosomes [6].

Mutation randomly changes one or more genes in a certain solution to increase diversity and avoid local optima [6]. It is important not to change too many genes in one solution, as this may render it unusable. Therefore, the mutation is often limited to the change of only one random activity in the schedule [6].

Mutation of all genes with a certain probability is used to get out of local optima and find more optimal solutions [6].

In the context of a scheduling problem, a population represents different schedule options. A crossover involves changes in class numbers, days, or classrooms between two different schedule variants for a particular group or instructor. Mutation makes similar random changes in an individual schedule.

In the genetic algorithm, selection is the process of selecting the best solutions for forming the next population. The essence of selection is to keep the best individuals and get rid of the worst ones, which leads to a gradual improvement of decisions in each subsequent iteration. The following selection methods were implemented [6]:

- *tournament*: two decisions are randomly chosen, priority is given to the decision with a higher fitness value;
- *ranking*: a selection method based on their suitability ranking;
- *roulette*: selection of candidates of the next generation based on probabilities corresponding to their fitness.

An individual's fitness determines his or her ability to effectively solve a task or meet certain optimality criteria. It is a numerical value that reflects how well a particular individual meets the requirements.

In the case of scheduling, an individual's fitness can be measured by the number of conflicts in the schedule, the efficiency of resource use, and the satisfaction of students' and teachers' needs. In its algorithm, the function takes into account such strict constraints as overlapping class schedules, inconsistencies with the recommended schedule, and the number of breaks between classes.

Scaling of the adaptability of individuals was implemented in selection. Scaling of fitness values is one of the optimization methods of genetic algorithms, which allows to increase the speed of convergence, unit/s of the algorithm. The basic idea is to rescale the fitness values so that they lie in the range from a to b. This can be done using a scaling function that transforms the original range of fitness values into a new range corresponding to the range from a to b [7].

Scaling of the fitness of individuals helps to reduce the influence of different scales of fitness functions on the results of the genetic algorithm and allows to more efficiently find the optimal solution [7].

Also, to increase the chances of finding an optimal solution, you can use elitism, which preserves a part of the best individuals from one generation to another without changes. The preservation of elite individuals from the previous population allows to preserve diversity and prevent the loss of useful information [8].

Criteria for stopping the algorithm:

- if the maximum number of iterations is reached;
- if a schedule is found, the fitness value of which is equal to 0.

One of the methods for improving the classical genetic algorithm is its island model. The GA island model is a model in which the population is divided into several groups (islands). Each island contains its own subpopulation that evolves independently of other islands [9].

With the help of the migration mechanism, individuals can move from one island to another to speed up the convergence of the algorithm. This will help reduce the risk of falling into local minima [9].

In the island model, subpopulations can use different algorithm parameters, different crossing, mutation and selection operators [9].

Also, the genetic algorithm, like its island model, can be speeded up by parallel implementation. Parallel implementation of GA will speed up the convergence time of the algorithm using several processors [10].

5. Software implementation of the developed modification

The main idea of the genetic algorithm is to gradually create and improve schedules using operators such as mutation, crossover, and selection. The basic steps of a genetic algorithm include initialization, main cycle, crossover, mutation, calculation of fitness values, selection, termination, and saving the results.

The island model of the genetic algorithm is similar to the classical genetic algorithm, but involves the creation of separate groups of populations that function in parallel and can interact by exchanging individuals.

The island model of the genetic algorithm includes such operations as the initialization of individual islands with different parameters and the possibility of migration of individuals between these islands.

The main steps of the GA island model include initialization, initialization of schedules, main loop, island traversal, crossover, mutation, calculation of fitness values, selection, migration (if necessary), selection of best individuals, search of best schedules, replacement of individuals, termination, search of best individual, completing the algorithm and saving the results.

6. Experiments and results

Input data for the system are:

- *data on classrooms*: classroom name, type, capacity, established departments;
- *data on departments*: name of the department and its abbreviated name;
- *data on specialties*: specialty name, code, department;
- *data on disciplines*: name, specialty, semester in which the discipline is taught;
- *group data*: group name, number of students, current semester, specialty;
- *data on teachers*: full name of the teacher and the department where he works;
- *data about classes*: discipline, class type, number of classes per week, teachers, groups, recommended audiences, recommended time.

The initial data is the optimal schedule of classes compiled for the selected department, which takes into account the recommendations for conducting and minimizes the following parameters:

- the number of overlapping classes for groups, teachers and classrooms;
- the number of windows between classes for groups and teachers.

A computer with the following characteristics was used for the tests: processor AMD Ryzen 5 2600 with a frequency of 3.9 GHz, 6 cores and 6 threads, the amount of RAM is 8.00 GB with a speed of 3200 MHz, the type of hard disk is SSD, and the operating system is the Virtual Machine Windows Server 2022.

The speed of convergence parameter was used to compare the performance of the algorithm. The speed of convergence shows how quickly or efficiently the algorithm approaches its optimal solution. This parameter can be calculated using the following formula (6):

$$S = \frac{|F_s - F_f|}{|t_f - t_s|} \tag{6}$$

where S is the speed of convergence; F_s is the initial fitness value;$F_(f)$ is the final value of fitness; t_s is the initial time of the algorithm; t_f is the final time.

The following abbreviations should also be noted: P_e is the value of elitism, i.e. the percentage of the total population; T_c is type of crossing: 1 – "custom one gene", 2 – "k-point"; N_k is the number of k crossing points; P_c is the probability of crossing; T_m is mutation type: 1 – "custom one gene", 2 – "all genes"; P_{mg} is probability of gene mutation; P_m is probability of mutation; T_s is selection type: 1 – "roulette", 2 – "ranging", 3 – "tournament"; T_s is initialization type: 1 – "random", 2 – "simple algorithm"; N_g is the maximum number of iterations of the algorithm; N_p is the size of the population.

The following parameters were used for the tests: number of days for making a schedule – 6, number of classes for making a schedule – 6, penalty for windows in groups – 1, penalty for windows in teachers – 1, penalty for overlapping classes – 5, and penalty for inconsistency recommended class time – 5.

Table 1 lists the test run parameters for the genetic algorithm, including the population size (500) and the maximum number of iterations (2000). According to table 1, the results of these tests are given. It is important to note that each entry in table 2 represents the average value of the results of three conducted experiments with the same parameters.

Genetic algorithm parameters, such as population size and maximum number of iterations, were kept constant for all tests to ensure comparability of results between different experiments.

Table 3 presents the parameters used during tests of the genetic algorithm modification – the island model. Table 4 contains the results of these tests, and each test was performed three times, after which the average value of the results was calculated. All tests used the same parameters: a population size of 500 and a maximum number of iterations of 2000.

The following designations should be entered: N_i is the number of islands; N_{st} is a step of increasing the values in the islands; N_{it} is the number of iterations through which to migrate between islands; P_{mig} is the number individuals participated in migration.

7. Discussion of research results

Analyzing table 2, we can draw the following conclusions about the highest convergence rate for the GA parameters: mutation probability – 0.2; mutation type – single gene mutation; crossover probability – 0.6; elitism value – 0.2; sampling type – roulette; crossover type – single gene crossover.

Table 1
Parameters for running genetic algorithm tests.

№	P_e	T_c	N_k	P_c	T_m	P_{mg}	P_m	T_s	T_i
1	0.1	1	-	0.7	1	-	0.2	1	1
2	0.1	1	-	0.7	1	-	0.4	1	1
3	0.1	1	-	0.7	1	-	0.6	1	1
4	0.1	1	-	0.7	2	0.05	0.2	1	1
5	0.1	1	-	0.7	2	0.1	0.2	1	1
6	0.1	1	-	0.7	2	0.2	0.2	1	1
7	0.1	1	-	0.4	1	-	0.2	1	1
8	0.1	1	-	0.6	1	-	0.2	1	1
9	0.1	1	-	0.8	1	-	0.2	1	1
10	0.2	1	-	0.6	1	-	0.2	1	1
11	0.3	1	-	0.6	1	-	0.2	1	1
12	0.1	1	-	0.6	1	-	0.2	2	1
13	0.1	1	-	0.6	1	-	0.2	3	1
14	0.1	2	2	0.6	1	-	0.2	3	1
15	0.1	2	4	0.6	1	-	0.2	3	1
16	0.1	2	6	0.6	1	-	0.2	3	1
17	0.1	1	-	0.6	1	-	0.2	3	2
18	0.1	1	-	0.6	1	-	0.4	3	2
19	0.1	1	-	0.8	1	-	0.2	3	2
20	0.1	1	-	0.6	2	0.1	0.2	3	2
21	0.3	1	-	0.6	1	-	0.2	3	2
22	0.1	2	4	0.6	1	-	0.2	3	2

To analyze the effect of the parameter $T_i = 2$ (using a modified initialization parameter), we calculate how much the algorithm's running time has decreased on average compared to the values of $T_i = 1$. To do this, let's calculate the average value of the algorithm execution time for both variants.

The average value of the GA running time at $T_i = 1$ is 1888.429 s.

The average value of the GA running time at $T_i = 2$ is 1115.938 s.

The reduction in the running time of the GA algorithm with $T_i = 2$ compared to $T_i = 1$ is approximately 40.9%. A 40.9% reduction in running time indicates an improvement in algorithm performance by using $T_i = 2$.

After analyzing table 4, we can draw the following conclusions about the fastest convergence rate for the island model parameters: number of islands – 6 (number of logical processor cores); parameter divergence step – 7; number of iterations to perform migration – 10; number of individuals to participate in migration – 0.1.

The average running time of the island model at $T_i = 1$ is 2335.027 s.

The average running time of the island model at $T_i = 2$ is 1313.531 s.

When using the modified initialization method, the island GA model is on average 1.8 times faster. This can improve the performance of the algorithm, which saves time and resources in solving the problem.

Table 2
Genetic algorithm results.

№	t_s, s	t_f, s	F_s	F_f	S
1	0.983	1937.365	1128.667	22.000	0.577
2	0.968	1852.033	1157.333	24.667	0.613
3	0.981	1884.862	1148.000	119.000	0.546
4	0.991	1840.709	1178.667	75.000	0.600
5	0.964	1865.484	1164.333	80.333	0.581
6	1.008	1881.805	1156.667	54.333	0.586
7	0.953	1851.585	1183.667	22.000	0.628
8	0.952	1796.640	1165.000	17.667	0.639
9	0.966	1808.654	1164.000	15.000	0.636
10	0.960	1794.841	1158.333	14.000	0.638
11	0.979	1840.432	1148.333	18.000	0.615
12	0.983	1878.421	1165.333	10.000	0.615
13	1.003	1839.763	1155.333	14.667	0.620
14	1.023	1980.765	1159.000	15.667	0.578
15	1.135	2196.617	1156.667	10.000	0.522
16	1.023	1980.765	1159.000	15.667	0.578
17	1.116	947.481	8.333	1.333	0.014
18	0.922	1334.211	9.667	1.333	0.008
19	0.914	1320.637	9.667	2.000	0.009
20	0.929	680.211	10.333	2.667	0.076
21	1.021	859.483	14.667	1.000	0.019
22	1.075	1559.583	6.333	1.333	0.016

Let's test the algorithms with the parameters that give the best results (table 5).

8. Conclusion

A modified genetic method has been developed that uses an initialization operator based on a priori information about the learning process that is available from the given constraints. The use of the developed approach to initializing the genetic method can significantly (several times) reduce the search time.

Also, a modified island model of the developed genetic method was developed to solve the problem of drawing up an optimal schedule of classes. The fundamental difference between the proposed method and existing analogues is the use of a modified initialization operator that tries to reduce the initial fitness value by simply searching through possible options. Using the modified initialization operator, the running time of the classical genetic algorithm was reduced by 40.9%. It is also worth noting that the running time of the island model of the genetic algorithm with the modified initialization operator was reduced by 1.8 times compared to the use of the classical initialization method. This means that the application of this method has significantly saved the algorithm execution time. The modified initialization method improves the performance and speed of the genetic algorithm for scheduling classes, which is important

Table 3
Parameters for launching genetic algorithm modification tests – island model.

№	P_e	T_c	P_c	T_m	P_m	T_s	T_i	N_i	N_{st}	N_{it}	P_{mig}
1	0.1	1	0.4	1	0.2	1	1	6	2	10	0.1
2	0.1	1	0.4	1	0.2	1	1	12	2	10	0.1
3	0.1	1	0.4	1	0.2	1	1	18	2	10	0.1
4	0.1	1	0.4	1	0.2	1	1	6	3	10	0.1
5	0.1	1	0.4	1	0.2	1	1	6	5	10	0.1
6	0.1	1	0.4	1	0.2	1	1	6	7	10	0.1
7	0.1	1	0.4	1	0.2	1	1	6	7	5	0.1
8	0.1	1	0.4	1	0.2	1	1	6	7	8	0.1
9	0.1	1	0.4	1	0.2	1	1	6	7	13	0.1
10	0.1	1	0.4	1	0.2	1	1	6	7	5	0.05
11	0.1	1	0.4	1	0.2	1	1	6	7	5	0.15
12	0.1	1	0.4	1	0.2	1	1	6	7	5	0.25
13	0.1	1	0.4	1	0.2	1	2	12	5	5	0.05
14	0.1	1	0.4	1	0.4	1	2	12	5	5	0.05
15	0.1	1	0.6	1	0.4	1	2	12	5	5	0.05
16	0.3	1	0.6	1	0.4	1	2	12	5	5	0.05

Table 4
Results of tests of modification of the genetic algorithm – island model.

№	t_s, s	t_f, s	F_s	F_f	S
1	1.144	1924.096	1123.667	14.667	0.577
2	2.211	3663.413	1101.000	6.333	0.299
3	3.068	5399.514	1118.667	4.000	0.207
4	1.123	1835.984	1119.333	9.333	0.605
5	1.121	1892.263	1112.667	6.000	0.585
6	1.100	1893.736	1094.333	5.333	0.576
7	1.160	1898.343	1131.000	5.000	0.593
8	1.060	1929.846	1124.000	8.000	0.579
9	1.120	1880.434	1152.333	10.667	0.608
10	1.070	1899.741	1143.000	5.667	0.599
11	1.080	1913.294	1109.000	7.000	0.576
12	1.161	1906.083	1146.667	11.000	0.596
13	1.497	1314.078	383.000	3.333	0.217
14	1.640	1819.898	10.667	0.667	0.012
15	1.606	1485.881	7.000	1.000	0.011
16	1.971	640.981	9.667	0.000	0.015

for effective problem solving.

An experimental study of the proposed genetic methods was performed. The island model of the genetic algorithm proved to be more efficient both in terms of speed and quality of the solutions obtained. On average, the island model of the GA works much faster – the execution time was reduced by 41.3% on average, it has a better fitness function result, the average fitness

Table 5
Test results of algorithms with the best parameters.

Algorithm	№	t_s, s	t_f, s	F_s	F_f
Genetic algorithm	1	1.077	847.888	8.000	0.000
	2	1.020	1320.838	11.000	0.000
	3	1.178	1946.044	12.000	2.000
	Average	1.092	1371.590	10.333	0.667
Island model of GA	4	1.857	2016.956	3.000	0.000
	5	1.009	349.973	12.000	0.000
	6	1.048	50.780	4.000	0.000
	Average	1.305	805.903	6.333	0.000

value is 0, which means that the algorithm has found the ideal solution and makes it a more efficient algorithm for this problem.

Acknowledgments

The work was carried out with the support of the state budget research project of the state budget of the National University "Zaporozhzhia Polytechnic" "Intelligent methods and tools for diagnosing and predicting the state of complex objects" (state registration number 0122U000972).

References

[1] S. Lukas, A. Aribowo, M. Muchri, Genetic algorithm and heuristic search for solving timetable problem case study: Universitas Pelita Harapan timetable, in: 2009 Second International Conference on the Applications of Digital Information and Web Technologies, 2009, pp. 629–633. doi:10.1109/ICADIWT.2009.5273979.

[2] I. A. Abduljabbar, S. M. Abdullah, An Evolutionary Algorithm for Solving Academic Courses Timetable Scheduling Problem, Baghdad Science Journal 19 (2021) 399–408. doi:10.21123/bsj.2022.19.2.0399.

[3] H. Alghamdi, T. Alsubait, H. Alhakami, A. Baz, A Review of Optimization Algorithms for University Timetable Scheduling, Engineering, Technology & Applied Science Research 10 (2020) 6410–6417. doi:10.48084/etasr.3832.

[4] O. M. Haitan, The university course timetabling automation, Scientific notes of Taurida National V.I. Vernadsky University. Series: Technical Sciences 1 (2020) 58–66. doi:10.32838/2663-5941/2020.2-1/09.

[5] A. R. Petrosian, R. V. Petrosyan, I. A. Pilkevych, M. S. Graf, Efficient model of PID controller of unmanned aerial vehicle, Journal of Edge Computing 2 (2023) 104–124. doi:10.55056/jec.593.

[6] V. Y. Snityuk, O. M. Sipko, Technology of Evolutionary Formation of Timetables in Institutions of Higher Education, Yu.V. Picha FOP Publisher, 2022.

[7] I. Mulyava, The solution of the problem of automated formation of schedule of educational institutions with genetic algorithms, Mizhnarodnyi naukovyi zhurnal “Internauka” (2018) 77–83. URL: http://nbuv.gov.ua/UJRN/mnj_2018_9%281%29__20.

[8] V. V. Kysil, I. V. Drach, T. M. Kysil, Task of creation and optimization of learning schedule that supplements s trict and volatile requirements model, Scientific notes of Taurida National V.I. Vernadsky University. Series: Technical Sciences 6 (2019) 65–70. doi:10.32838/2663-5941/2019.6-1/12.

[9] W. A. Indha, N. S. Zamzam, A. Saptari, J. A. Alsayaydeh, N. b. Hassim, Development of Security System Using Motion Sensor Powered by RF Energy Harvesting, in: 2020 IEEE Student Conference on Research and Development (SCOReD), 2020, pp. 254–258. doi:10.1109/SCOReD50371.2020.9250984.

[10] JavaTpoint, Genetic Algorithm in Machine Learning, 2021. URL: https://www.javatpoint.com/genetic-algorithm-in-machine-learning.

Predictive machine learning of soybean oil epoxidizing reactions using artificial neural networks

Bogdan B. Sus[1], Oleksandr S. Bauzha[1], Sergiy P. Zagorodnyuk[1], Taras V. Chaikivskyi[2] and Oleksandr V. Hryshchuk[1]

[1]*Taras Shevchenko National University of Kyiv, 64/13 Volodymyrska Str., Kyiv, 01601, Ukraine*
[2]*Lviv Polytechnic National University, 12 Bandera Str., Lviv, 79013, Ukraine*

Abstract

Experimental data on the epoxidation process of soybean oil with hydrogen peroxide/acetic anhydride were analyzed. This study utilizes experimental data to construct a training dataset for neural network training. Post-training, the neural network facilitates the optimization of epoxy curing reaction parameters, monitors its evolution, and refines the epoxy product synthesis process. Furthermore, a novel methodology has been devised to calculate the outcomes of epoxyation in unsaturated compound mixtures. This method empowers precise control over the epoxyation process at the synthesis phase, under specific reaction conditions, and elevates the technology involved in epoxy output production.

Keywords

Neural network, Optimization, Soybean oil, Epoxidizing

1. Introduction

Modern information technologies enable the solution of a wide range of practical tasks that involve the procedure of deep machine learning and subsequent outcome testing [1]. At present, there are already many areas of human activity where diverse computational tasks have been successfully solved using artificial neural networks. For instance, by analyzing the electroencephalogram of a person expressing a particular phrase, using deep feedforward neural networks, it has been possible to build a human speech synthesizer [2]. Such a synthesizer is capable of guessing and articulating certain fixed expressions used by a person in the past but not recalled during the operation of the synthesizer. The use of neural networks in medicine is essentially a separate branch of science and technology: they can be used for interpreting X-ray images, making it easy to determine the degree of bone wear and tear in a person, as well as artificial implants, fasteners, and prosthetics injected into the patient's body [3].

CS&SE@SW 2023: 6th Workshop for Young Scientists in Computer Science & Software Engineering, February 2, 2024, Kryvyi Rih, Ukraine

bnsuse@gmail.com (B. B. Sus); asb@univ.kiev.ua (O. S. Bauzha); szagorodniuk@gmail.com (S. P. Zagorodnyuk); taras.v.chaikivskyi@lpnu.ua (T. V. Chaikivskyi); oleksandr_hryshchuk@knu.ua (O. V. Hryshchuk)
https://iht.knu.ua/staff/sus-b-b (B. B. Sus); http://ced.knu.ua/list/bauzha (O. S. Bauzha); http://ced.knu.ua/list/zagorodnyuk (S. P. Zagorodnyuk); https://pro100.lpnu.ua/moodle/user/profile.php?id=2&lang=en (T. V. Chaikivskyi)
0000-0002-2566-5530 (B. B. Sus); 0000-0002-4920-0631 (O. S. Bauzha); 0000-0003-3415-7746 (S. P. Zagorodnyuk); 0000-0003-3415-7746 (T. V. Chaikivskyi); 0009-0007-9926-4231 (O. V. Hryshchuk)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

CEUR-WS.org/Vol-3662/paper08.pdf

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

Chemists and biologists use neural networks to predict the outcomes of chemical and biological processes, select optimal conditions for such processes, and assess the quantities and types of catalysts [4] or inhibitors [5]. In the field of robotics and automation, neural networks are employed for the approximate calculation and prediction of motion trajectories for automated machines, mechanisms, and manipulators [6], as well as for optimizing the consumption of energy resources and materials. Among the natural sciences, including mathematics, mechanics, materials science, vibrations and waves, spectral analysis, the application of neural networks enables the solution of classical fundamental equations, including those involving Bessel and Neumann special functions [7], which describe real multi-dimensional systems. Numerical solutions to such equations were previously inaccessible.

Neural network structures are utilized in various branches and directions of physics. This includes mathematical modeling of the output of photovoltaic panels and approximating generated power [8], processing and analysis of sensor data and measurement devices, and substance identification.

Neural networks find extensive applications in physical problems related to the propagation, interference, diffraction, and absorption of electromagnetic waves [9, 10], as well as in ultrasound for biological research [11] and for the study of massive solid media [12].

The main goal and principle of green chemistry are the utilization of renewable, environmentally friendly raw materials, which will contribute to reducing biodegradation and the toxicity of industrial production [13]. The polymerization of soybean oil is essential for the creation of polymers used in the production of printing and textile inks. Epoxydized oils possess unique chemical properties and, therefore, have a wide range of applications. They are used to enhance the operational quality of rubber, as components for producing photographic films, packaging materials in medicine, and in the production of food products.

There is an urgent need for a fast and accurate method for identifying the true content of oil mixtures. In this study, Raman spectroscopy, combined with three deep learning models (CNN-LSTM, enhanced AlexNet, and ResNet), was used for the simultaneous determination of the quantities of extra virgin olive oil (EVOO), soybean oil, and sunflower oil in a blend of olive oils. The research demonstrated that all three deep learning models outperformed traditional chemometric methods in predicting the composition [14]. Currently, active research is being conducted on the synthesis of solid polymer materials using soybean oil with mechanical properties that can be utilized as construction and building materials. Epoxides belong to the group of cyclic ethers - metabolites that are often formed by cytochromes, acting on aromatic or double bonds. The specific site on a molecule that undergoes epoxidation is called the site of epoxidation (SOE). Thus, artificial intelligence methods significantly enhance the accuracy of SOE molecule identification and the selection of optimal parameters for controlling the epoxidation process. Indeed, this type of learning is facilitated by artificial neural networks (ANN), which are essentially a technical software implementation of the biological neural structure in the human brain. Such neural networks are also commonly referred to as connection systems because, for such ANNs, a fundamental component is the system of connections and links that enable one neuron to establish pathways for the propagation of signals to other neurons and, conversely, receive similar signals from them. Epoxy oils are widely used in the production of polyvinyl chloride (PVC) and polymers based on it because epoxy oils are among the best stabilizers and plasticizers for such polymers. Epoxy oils offer specific advantages

over stabilizers of other types. For example, adding epoxy stabilizers to a polymer significantly increases its thermal stability and provides resistance to ionizing radiation. Epoxy oils can be used as curing agents and binders (in compositions with various oligomers) as well as stabilizers (in PVC compositions). The production and use of epoxy oils continue to grow steadily as they are integral components of paint and varnish products based on epoxy, cellulose ether oligomers, PVC, and plasticizers for organodispersed coatings.

In this study, a deep learning algorithm is employed for the automatic detection of issues in pipelines containing epoxy oils [15]. In [16] the construction of an electromagnetic field shielding system using a multi-layer perceptron neural network-based system designed for predicting electromagnetic absorption by composite films based on polycarbonate and carbon nanotubes is examined. The proposed system includes 15 different multi-layer perceptron networks [17].

Organic peracids are utilized for the implementation of liquid-phase epoxidation of unsaturated organic substances [18].

2. Experimental

Nykulyshyn et al. [19] suggests the use of an epoxidation system consisting of H_2O_2/acetic acid/catalyst. In this system, the epoxidizing agent is also an organic peracid formed through the interaction of H_2O_2 with organic acid in the presence of a catalyst (figure 1). In this case, the organic acid circulates within the system, with its molecule periodically gaining an additional oxygen atom, which is then transferred to another substance.

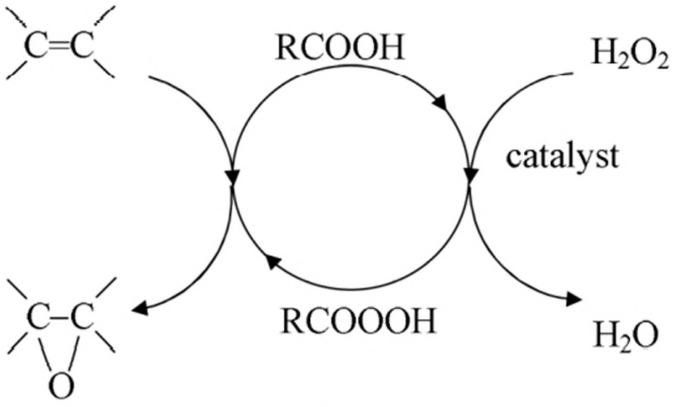


Figure 1: Epoxidizing of the substance using the H_2O_2/organic acid/catalyst system.

The authors propose using the anhydride of organic acid instead of organic acid, which will reduce the formation of water in the reaction mixture and accelerate the reaction of nadoctovoic acid formation.

The objective of this study is to improve and modify the technology for obtaining epoxy oils, establish optimal conditions for the economically viable production of epoxy oils that meet quality standards (figure 1).

Table 1
Quality indicators of epoxidized oils.

Physical-chemical indicator	Norm for brands (technical conditions)*		
	ST	SU	C
Epoxy number, % (oxyran oxygen content), not less than	6.5	6.4	6.0
Iodine number, g I2/100 g, no more than	1.5	2.0	8.0

* As stabilizers and plasticizers for PVC-based polymers

The practical value of using the process model obtained through neural networks lies in its ability to monitor the quality of epoxy oil during the synthesis stage. The dependencies of the time parameters of the chemical epoxidation reaction on:

- Initial concentration of acetic anhydride
- Initial concentration of hydrogen peroxide
- Initial concentration of ion-exchange resin KU-2x8
- Process temperature.

The obtained dependencies allow researchers the assessment of the optimal concentration of epoxying mixture reagents, process duration, and temperature.

Experimental data indicates that the relationship between the epoxy number of epoxidized soybean oil and temperature at various concentrations of acetic anhydride, hydrogen peroxide, catalyst amount, and process duration exhibits a complex nature. In the initial stages, the reaction rate steadily increases with rising temperature and catalyst concentration. However, at excessively high temperatures (>348 K) and catalyst concentrations (>15 g per 100 cm^3), there is a possibility of a decrease in the achieved epoxy number due to secondary reactions involving the opening of epoxy cycles.

Therefore, to determine the optimal conditions for the chemical process, it is advisable to construct a mathematical model and apply it to calculate the parameters of such processes. The following assumptions were made during the model creation:

- X1 – concentration of acetic anhydride, wt.%, 2<X1<9
- X2 – concentration of hydrogen peroxide 46%, wt.%, 25<X2<40
- X3 – amount of catalyst, wt.%, x3<15
- X4 – temperature, K, 333<X4<353
- X5 – duration of the process, min, X5<360

3. Methods

For training a neural network using the results of experiments, a target dataset has been created. The neural network's inputs characterize the experimental conditions, while the outputs characterize the final results, specifically the concentrations of reaction products.

1. The input parameters of the neural network are as follows:

2. Concentration of Acetic Anhydride (Normalized)
3. Concentration of Hydrogen Peroxide (Normalized)
4. Catalyst KU-2×8 (Kat) (Normalized)
5. Normalized Temperature
6. Normalized Reaction Time
7. Initial Epoxy Number (Normalized)
8. Initial Iodine Number (Normalized)

The calculations were made with the assumption that the initial values of the Epoxy and Iodine numbers influence the course of the experiment. The experimental data sample did not exceed 400 minutes, so normalization was performed based on this time. The reactions were conducted in a temperature range of 423 to 443 K. The output parameters of the neural network are as follows:

1. Final Epoxy Number (Normalized)
2. Final Iodine Number (Normalized)

In figure 2, a five-layer neural network is depicted. It was trained using input experimental data obtained during the epoxidation of soybean oil. Seven input parameters were selected for the neural network, and the values of each input parameter were normalized to 1. The neural network's output layer consists of 2 neurons. This presented neural network includes three hidden layers, each of which contains 20 neurons.

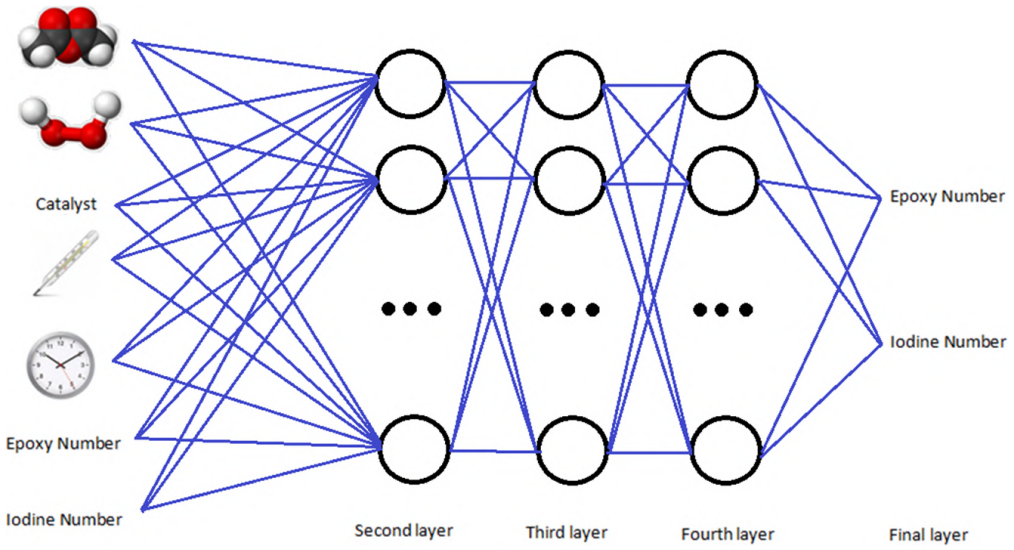


Figure 2: A five-layer neural network.

4. Results and discussion

As observed in figure 3, which was generated after training the neural network, an increase in reaction temperature leads to a rapid rise in the Epoxy Number within the first 100 minutes of

the reaction. Further progression of the reaction results in a plateau of Epoxy Number values, and continued reaction may even lead to a significant decrease.

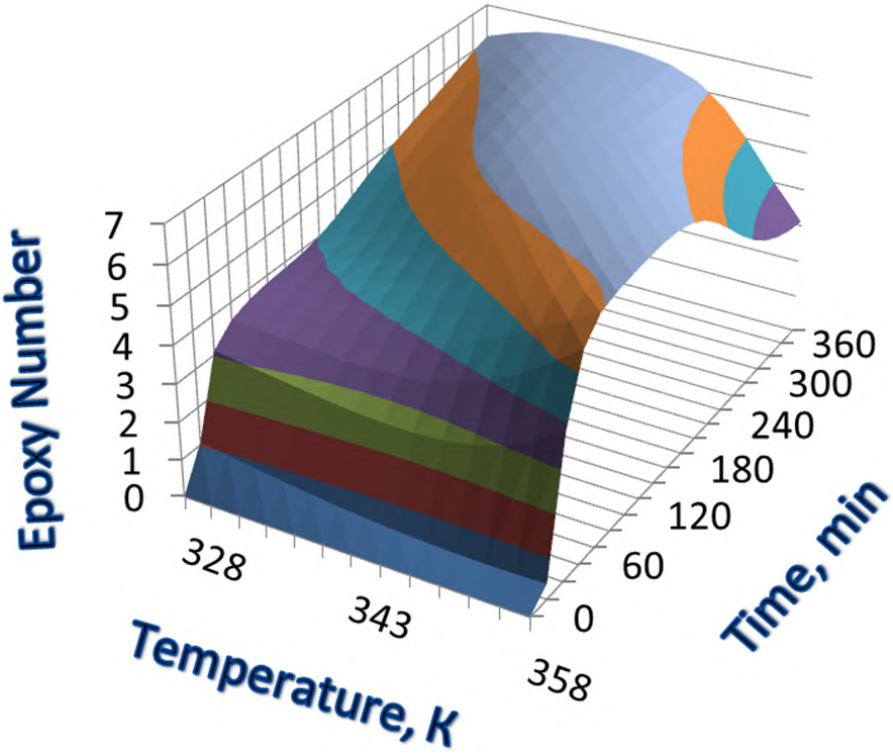


Figure 3: The relationship between the Epoxy Number and the reaction temperature and reaction time during the epoxidation of soybean oil. The calculations were conducted while interacting a soybean oil solution in toluene with the epoxidation system consisting of H_2O_2, acetic anhydride, and a catalyst. The concentrations used were as follows: Acetic anhydride concentration = 5.72 wt.%, Hydrogen peroxide concentration = 33.1 wt.%, and catalyst concentration = 5 wt.%.

The training of the neural network is accomplished using a training mechanism. To find the minimum of the error function, the backpropagation algorithm is applied, utilizing the stochastic gradient descent method.

At lower reaction temperatures, within the initial 60 minutes, the Epoxy Number increases rapidly to values of 3-4, with subsequent growth occurring more slowly. The data obtained during the experiment are limited in both quantity and the range of argument values. However, through neural network training, the system extrapolates the obtained values to create a more detailed graph by extending the range of argument values.

Figure 4 provides a comparison between experimental data and data obtained after training the neural network, which covers a wider range. Experimental data is represented by the red lines, while the predictions made by the neural network are shown in blue lines. The calculations were conducted during the interaction of a soybean oil solution in toluene with the epoxidation system involving H_2O_2, acetic anhydride, and a catalyst. The concentrations used were as follows: Acetic Anhydride concentration = 5.72 wt.%, Hydrogen Peroxide concentration = 33.1

wt.%, and catalyst concentration = 5 wt.%. The reaction temperatures are as follows: a – T=333 K, b – T=343 K, c – T=353 K. Through training, the neural network accurately processes the input data and predicts experimental outcomes beyond the scope of the available experimental data.

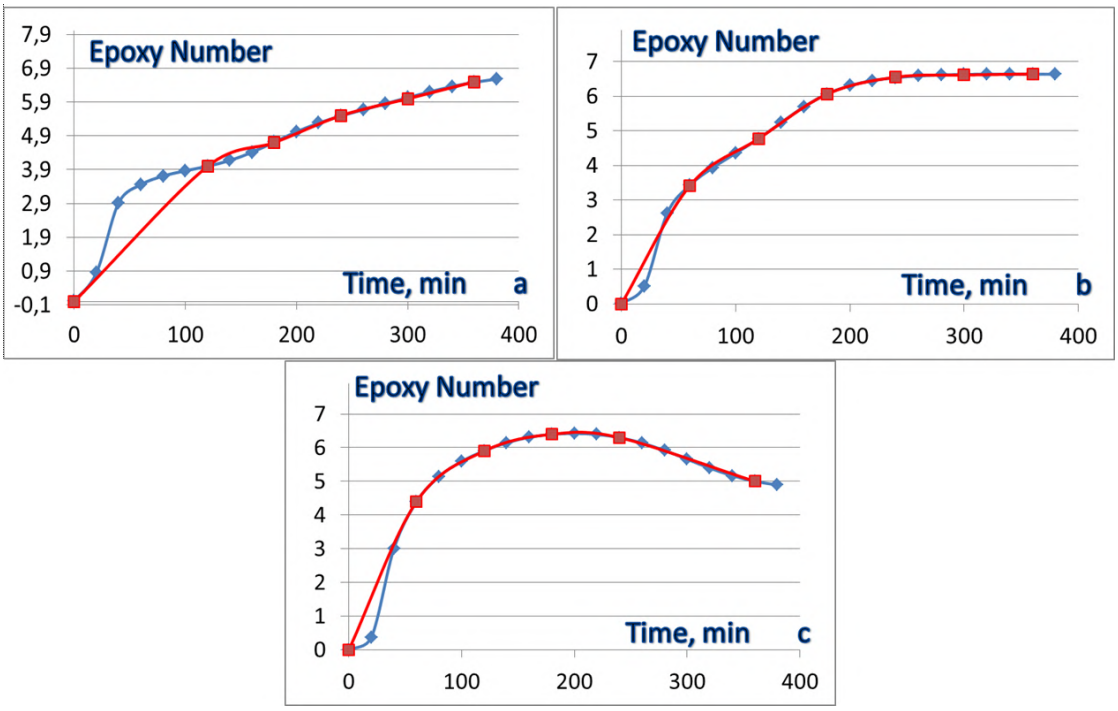


Figure 4: Comparison of Epoxy Number dependencies on reaction temperature and reaction time during the epoxidation of soybean oil. The reaction temperatures are as follows: a – T=333 K, b – T=343 K, c – T=353 K

The convexity observed in the dependency, as shown in figure 4a, at small reaction times is a result of the network being trained on intermediate experiment results. In other words, the network considered the entirety of the input data.

As seen in figure 5, which was generated after training the neural network, an increase in reaction temperature leads to a rapid decrease in the Iodine Number. The calculations were carried out during the interaction of a soybean oil solution in toluene with the epoxidation system comprising H_2O_2, acetic anhydride, and a catalyst. The concentrations used were as follows: Acetic anhydride concentration = 5.72 wt.%, Hydrogen peroxide concentration = 33.1 wt.%, and catalyst concentration = 5 wt.%. However, with a longer reaction time, there is a slight rebound in the Iodine Number towards higher values. At lower reaction temperatures in the mixture, the Iodine Number exhibits a much slower decreasing trend.

Similar to the data in figure 5, the data obtained during the experiment are limited both in quantity and the range of argument values. Nevertheless, through training, the neural network expands the range of argument values and interpolates the obtained values to create a more detailed graph.

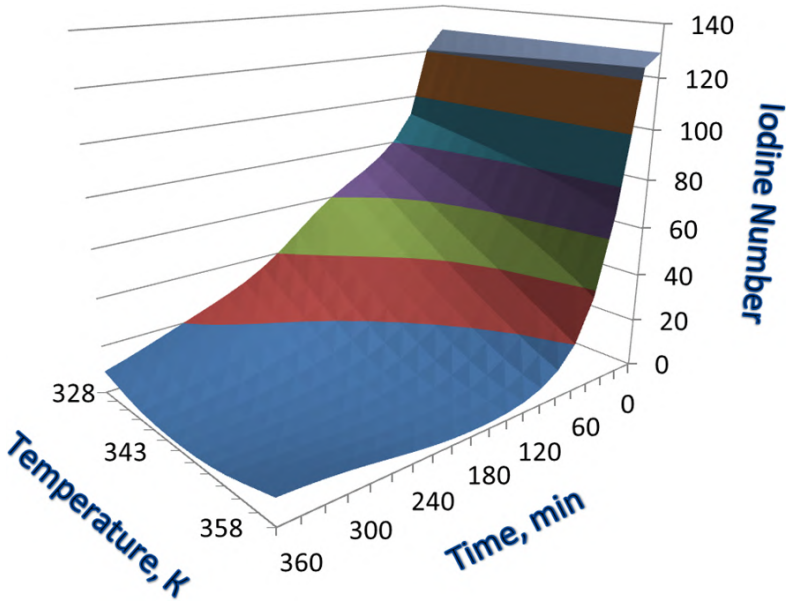


Figure 5: Dependency of the Iodine Number on the reaction temperature and reaction time during the epoxidation of soybean oil.

In figure 6, a comparison between experimental data and data obtained after neural network training is presented. As a result of training, the neural network accurately reproduces the experimental data and predicts experiment outcomes beyond it.

The wave-like dependencies observed in figure 5a at small reaction times and the convex shapes in dependencies figure 6(b,c) are predicted by the neural network. The first experimental data point in figure 5a was taken after 120 minutes, and the true dependence of the Iodine Number on time was unknown. However, based on the entire array of experimental data, the neural network predicts that the Iodine Number should vary according to this pattern.

After training, the neural network operates with output functions of multiple variables (in our case, 7 variables) and allows us to examine the projections of calculated (learned) functions onto selected parameter axes that are of interest to us. For example, in figure 7a, the relationship between the Epoxy Number after two hours of the epoxidation reaction of soybean oil and the concentration of Acetic Anhydride (OA) and temperature is shown. The reaction time, the concentration of H_2O_2, and the catalyst concentration are held constant.

Simultaneously reducing the concentration of OA and increasing the reaction temperature leads to the highest Epoxy Number value (figure 7a). Furthermore, at high OA concentrations and low reaction temperatures, a local maximum in the Epoxy Number dependence is observed within the selected range. The reaction performs the least efficiently at moderate OA concentrations and low reaction temperatures. It is important to note that in this case, we have considered a limited reaction time (time = 120 minutes). With a different reaction time, the dependencies may vary, and the optimal parameters (OA concentration and temperature) may fall within different ranges.

In figure 7b, we observe similar dependencies to those in figure 6, but for the Iodine Number.

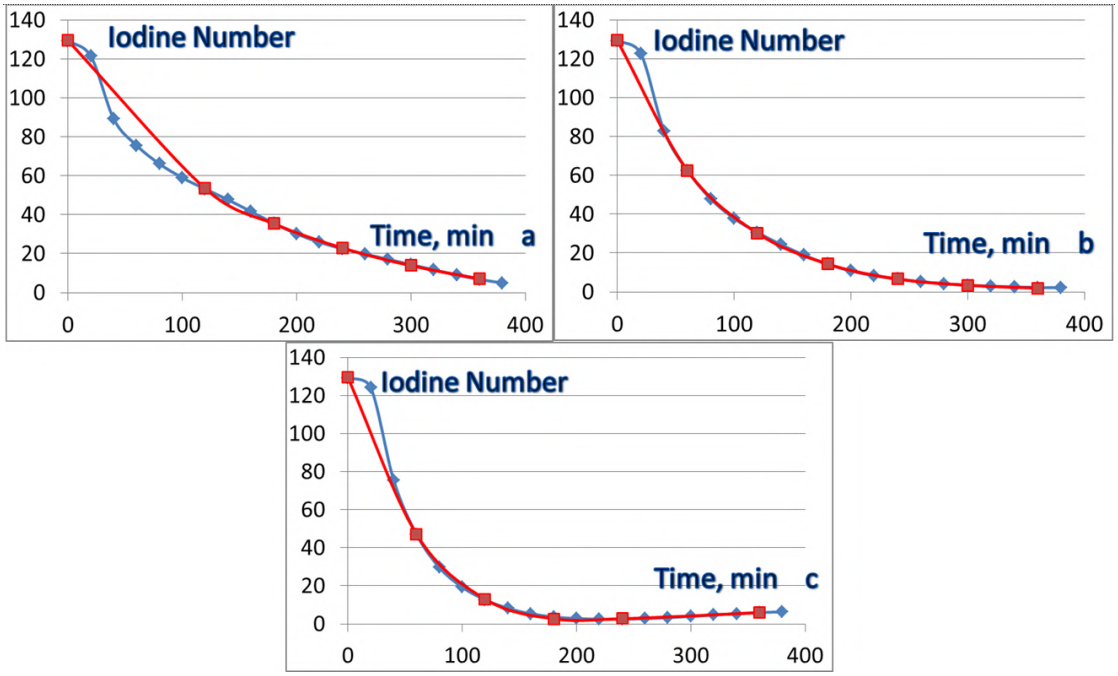


Figure 6: Comparison of Iodine Number dependencies on reaction temperature and reaction time during the epoxidation of soybean oil. Experimental data is represented by the red lines, while the predictions made by the neural network are shown as blue lines. The calculations were conducted during the interaction of a soybean oil solution in toluene with the epoxidation system consisting of H_2O_2, acetic anhydride, and a catalyst. The concentrations used were as follows: Acetic Anhydride concentration = 5.72 wt.%, Hydrogen Peroxide concentration = 33.1 wt.%, and catalyst concentration = 5 wt.%. Reaction temperatures are denoted as follows: a – T=333 K, b – T=343 K, c – T=353 K

As seen in the figure, even at the maximum reaction temperature (within the selected range) and the highest OA concentration, the Iodine Number does not reach its minimum value.

Optimal concentrations of OA and reaction temperature were calculated to achieve the maximum Epoxy Number and the minimum Iodine Number during the interaction of a soybean oil solution in toluene with the epoxidation system H_2O_2/acetic anhydride/catalyst (as shown in figure 7a and figure 7b). Conditions:

- X2 = 33.1 wt.% (Hydrogen Peroxide concentration)
- X3 = 5 wt.% (Catalyst concentration)
- X5 = 120 minutes (Reaction time)

Optimal Epoxy Number values occur at:

- X1 = 3.5 wt.% (Acetic Anhydride concentration)
- X4 = 358 K (Reaction temperature)

Optimal Iodine Number values occur at:

- X1 = 7.5 wt.% (Acetic Anhydride concentration)

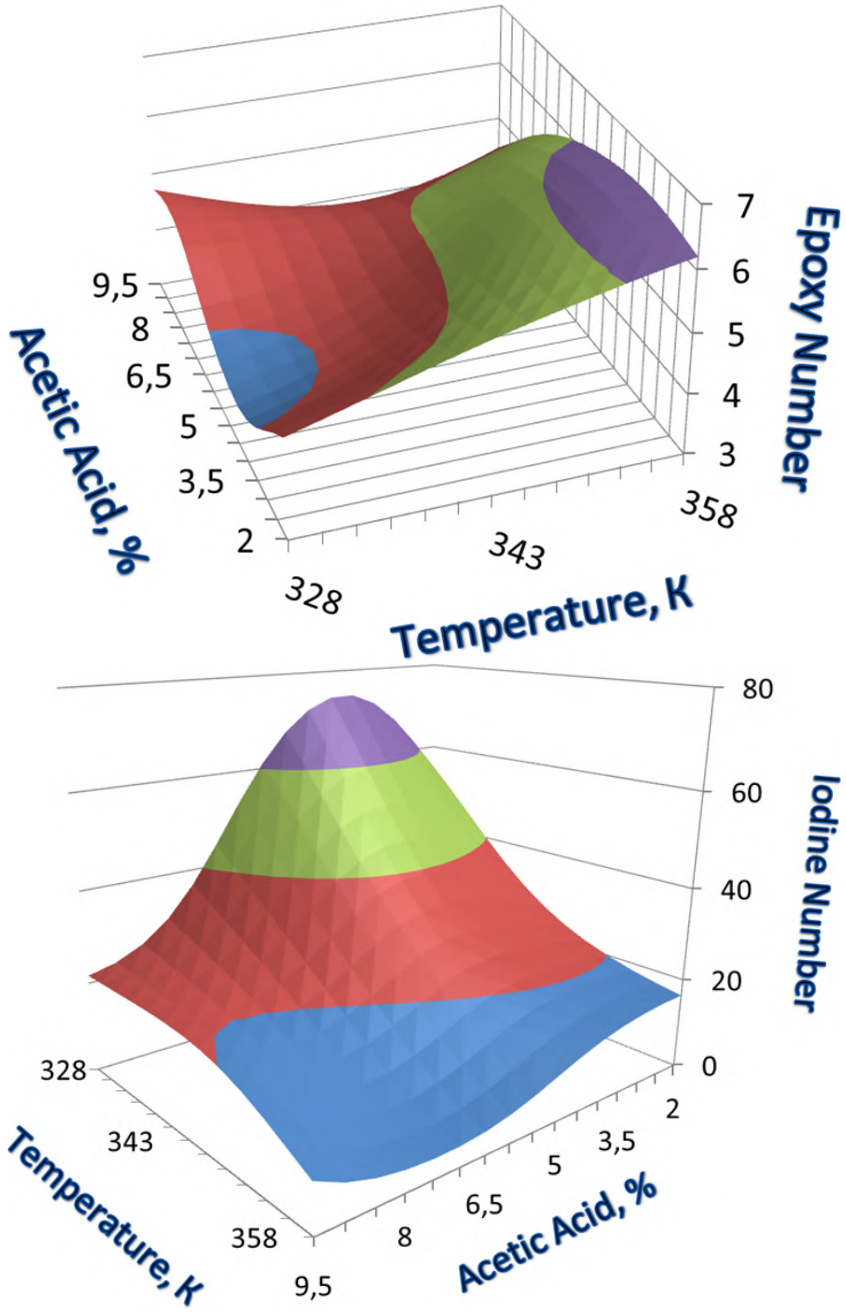


Figure 7: Dependency of Epoxy Number (a) Iodine Number (b) on the Acetic Anhydride concentration and temperature during the epoxidation of soybean oil. The calculations were carried out during the interaction of a soybean oil solution in toluene with the epoxidation system consisting of H_2O_2, acetic anhydride, and a catalyst. The catalyst concentration in the mixture was 5 wt.%, the Hydrogen Peroxide concentration was 33.1 wt%, and the reaction time was 120 minutes.

- X4 = 358 K (Reaction temperature)

As seen in figure 8a, the Epoxy Number reaches its maximum values at the highest concentration of H_2O_2 in the solution and the maximum reaction temperature (within the specified range of input values).

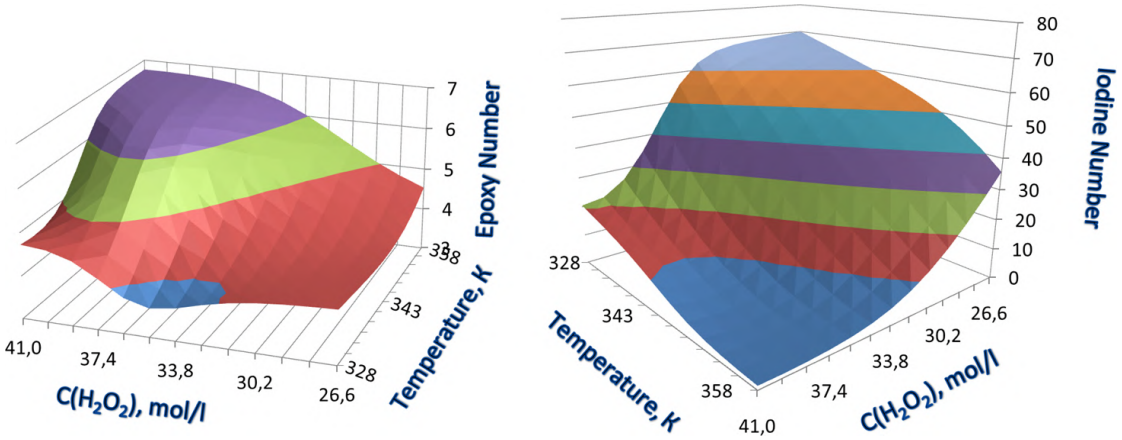


Figure 8: Dependency of the Epoxy Number (a) Iodine Number (b) on the concentration of H_2O_2 and reaction temperature during the epoxidation of soybean oil. The calculations were performed during the interaction of a soybean oil solution in toluene with the epoxidation system consisting of H_2O_2, acetic anhydride, and a catalyst. The mixture contained a catalyst concentration of 5 wt.%, Acetic Anhydride concentration of 5.72 wt.%, and a reaction time of 120 minutes

As observed in figure 8b, the Iodine Number reaches its minimum values at the highest concentration of H_2O_2 in the solution and the maximum reaction temperature (within the specified range of input values).

As observed in figure 9a, at the maximum reaction temperature, the maximum Epoxy Number is achieved after 2 hours of reaction across the entire specified range of catalyst concentration values.

As observed in figure 9b, at the maximum reaction temperature, the minimum Iodine Number is achieved after 2 hours of reaction across the entire specified range of catalyst concentration values.

In conclusion, it is important to emphasize that following training on a limited range of input experimental data, the neural network enables the expansion (prediction) of the range of experimental data. The addition of new data to the experimental database, which is used to train the neural network, offers the opportunity to refine the dependencies further.

5. Conclusions

The method presented in this study for determining the parameters of the soybean oil epoxidation process has demonstrated the nature of the dependence of these chemical reaction parameters on the concentration of OA and H_2O_2, the amount of catalyst, reaction temperature, and reaction time. The use of the neural network clearly illustrates the ability to quantitatively

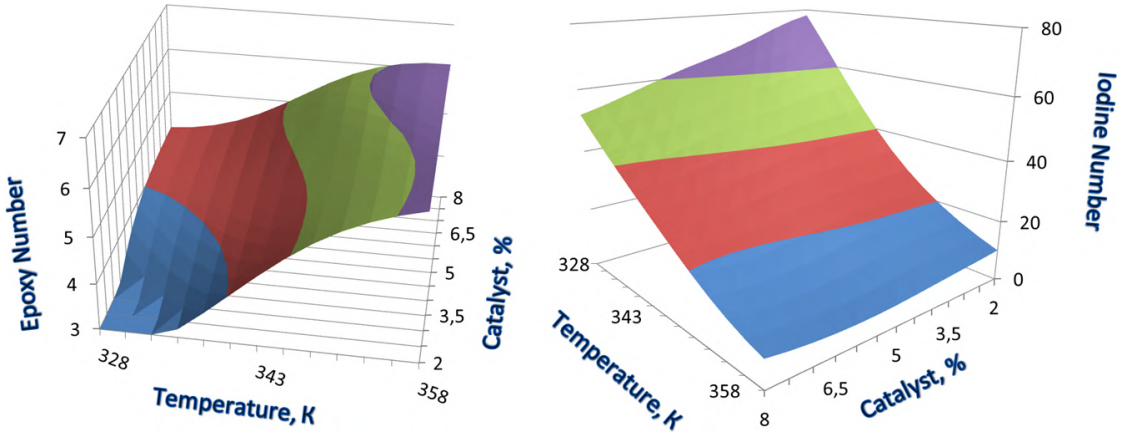


Figure 9: Dependency of the Epoxy Number (a) Iodine Number (b) on the catalyst concentration and reaction temperature during the epoxidation of soybean oil. The calculations were conducted during the interaction of a soybean oil solution in toluene with the epoxidation system involving H_2O_2, acetic anhydride, and a catalyst. The mixture had a Hydrogen Peroxide concentration of 33.1 wt.%, Acetic Anhydride concentration of 5.72 wt.%, and a reaction time of 120 minutes

predict the outcome of soybean oil epoxidation with a limited sample of experimental data. The prediction of results includes determining the epoxy number, iodine number, and additional information about the progress of the reaction.

The application of the trained neural network allows for the determination of optimal conditions for conducting the epoxidation process of oil using the epoxidation system, which consists of acetic anhydride, hydrogen peroxide, and the catalyst KU-2x8, with specific values for duration and temperature.

While experimental studies were conducted on soybean oil, the dependencies obtained in this work are also applicable to other vegetable oils such as castor oil, rapeseed oil, flaxseed oil, sunflower oil, and olive oil. Experiments on these vegetable oils are already underway.

Acknowledgments

This work has been supported by Ministry of Education and Science of Ukraine: Grant of the Ministry of Education and Science of Ukraine for perspective development of a scientific direction “Mathematical sciences and natural sciences” at Taras Shevchenko National University of Kyiv.

In recognition of Dr. Orest Pyrig’s substantial contributions to chemical research, and OJSC “Gallak” (Borislav) for their provision of scientific laboratory facilities, equipment, and reagents.

References

[1] B. R. Bartoldson, B. Kailkhura, D. Blalock, Compute-Efficient Deep Learning: Algorithmic Trends and Opportunities, 2022. doi:10.48550/ARXIV.2210.06640.

[2] M. Hernandez, N. Ventura-Campos, A. Costa, A. Miro-Padilla, C. Avila, Brain networks involved in accented speech processing, Brain and Language 194 (2019) 12–22. doi:10.1016/j.bandl.2019.03.003.

[3] G. Urban, S. Porhemmat, M. Stark, B. Feeley, K. Okada, P. Baldi, Classifying shoulder implants in x-ray images using deep learning, Computational and Structural Biotechnology Journal 18 (2020) 967–972. doi:10.1016/j.csbj.2020.04.005.

[4] H. Li, Z. Zhang, Z. Liu, Application of Artificial Neural Networks for Catalysis: A Review, Catalysts 7 (2017) 306. doi:10.3390/catal7100306.

[5] S. C. Nayak, B. B. Misra, H. S. Behera, Artificial chemical reaction optimization of neural networks for efficient prediction of stock market indices, Ain Shams Engineering Journal 8 (2017) 371–390. doi:10.1016/j.asej.2015.07.015.

[6] J. Sweafford Jr., F. Farbod, A neural network model-based control method for a class of discrete-time nonlinear systems, Mechatronic Systems and Control 49 (2021). doi:10.2316/J.2021.201-0134.

[7] K. O. Lye, S. Mishra, D. Ray, Deep learning observables in computational fluid dynamics, Journal of Computational Physicsl 410 (2020) 109339. doi:10.1016/j.jcp.2020.109339.

[8] A. Saberian, H. Hizam, M. A. M. Radzi, M. Z. A. Ab Kadir, M. Mirzaei, Modelling and Prediction of Photovoltaic Power Output Using Artificial Neural Networks, International Journal of Photoenergy 2014 (2014) 1–10. doi:10.1155/2014/469701.

[9] S. Tani, Y. Kobayashi, Ultrafast laser ablation simulator using deep neural networks, Scientific Reports 12 (2022) 5837. doi:10.1038/s41598-022-09870-x.

[10] T. Kusumoto, K. Mori, Prediction of ultrashort pulse laser ablation processing using machine learning, in: Laser Applications in Microelectronic and Optoelectronic Manufacturing (LAMOM) XXVI, Online Only, United States, 2021, p. 1. doi:10.1117/12.2577447.

[11] H. Zhang, J. Lian, Z. Yi, R. Wu, X. Lu, P. Ma, Y. Ma, HAU-Net: Hybrid CNN-transformer for breast ultrasound image segmentation, Biomedical Signal Processing and Control 87 (2024) 105427. doi:10.1016/j.bspc.2023.105427.

[12] K. Shukla, P. C. Di Leoni, J. Blackshire, D. Sparkman, G. E. Karniadakis, Physics-Informed Neural Network for Ultrasound Nondestructive Quantification of Surface Breaking Cracks, Journal of Nondestructive Evaluation 39 (2020) 61. doi:10.1007/s10921-020-00705-1.

[13] X. Jia, T. Wang, H. Zhu, Advancing Computational Toxicology by Interpretable Machine Learning, Environmental Science and Technology 57 (2023) 17690–17706. doi:10.1021/acs.est.3c00653.

[14] X. Wu, X. Zhang, Z. Du, D. Yang, B. Xu, Raman spectroscopy combined with multiple one-dimensional deep learning models for simultaneous quantification of multiple components in blended olive oil, Food Chemistry 431 (2024) 137109. doi:10.1016/j.foodchem.2023.137109.

[15] Y.-Y. Huang, H.-H. Wu, B.-H. Cai, Y.-C. Tang, J.-Y. Tseng, B.-H. Lu, AI Defect Recognition System for Epoxy Resin Pipeline in Factory, in: 2022 IEEE International Conference on Consumer Electronics - Taiwan, Taipei, Taiwan, 2022, pp. 69–70. doi:10.1109/ICCE-Taiwan55306.2022.9869229.

[16] L. Sidi Salah, M. Chouai, Y. Danlée, I. Huynen, N. Ouslimani, Simulation and Optimization of Electromagnetic Absorption of Polycarbonate/CNT Composites Using Machine

Learning, Micromachines 11 (2020) 778. doi:10.3390/mi11080778.
[17] J. Kantee, S. Kajorncheappunngam, Characterization of Epoxidized Rubber Seed Oil, Key Engineering Materials 728 (2017) 295–300. doi:10.4028/www.scientific.net/KEM.728.295.
[18] I. Nykulyshyn, T. Chaikivskyi, Z. Pikh, Research into epoxidation process of the c9 fraction hydrocarbons and unsaturated co-oligomers by peroxyethanoic acid, Eastern-European Journal of Enterprise Technologies 2 (2018) 53–60. doi:10.15587/1729-4061.2018.126682.
[19] I. Nykulyshyn, Z. Pikh, L. Shevchuk, S. Melnyk, Examining the epoxidation process of soybean oil by peracetic acid, Eastern-European Journal of Enterprise Technologies 3 (2017) 21–28. doi:10.15587/1729-4061.2017.99787.

Software development of thermal resistance calculator for thermal insulation parameters determines dielectric building structures

Vitalii M. Bazurin[1], Oleg I. Pursky[1], Yevheniya M. Karpenko[2], Tetiana V. Pidhorna[1] and Andrey I. Nechepourenko[1]

[1]State University of Trade and Economics, 19 Kyoto Str., Kyiv, 02156, Ukraine
[2]Ivan Franko Zhytomyr State University, 40 Velyka Berdychivska Str., Zhytomyr, 10008 , Ukraine

Abstract

The article presents the specifics of the software implementation of the thermal resistance calculator named "ThermoResist". The developed thermal resistance calculator calculates the thermal resistance of the wall, window, attic floor, and roof in accordance with the State Building Regulations of Ukraine. The programming language C♯ was chosen to implement computation procedures and the thermal resistance calculator interface. Calculations were performed under the assumption that the contributions of different thermal resistance mechanisms to the total thermal resistance of dielectric building materials are additive.

Keywords

thermal resistance, calculator, heat transfer processes, dielectric building structures

1. Introduction

Digitization of all spheres of public life and industry requires the development of appropriate tools [1]. One of the most important branches of industry in Ukraine is the construction industry because after the war it will be necessary to rebuild hundreds of destroyed cities and villages, to rebuild thousands of houses. Thermal insulation of buildings is an important part of construction. Both civil engineers and ordinary workers in the construction industry will carry out such work. The training of specialists in the construction industry is carried out by institutions of higher and professional pre−higher education. That is why the ability to perform thermal insulation calculations is an important component of the professional competence of civil engineers. The creation of thermal insulation begins with its calculations. Such calculations are usually carried out manually. At the same time, there are thermal resistance calculators that

CS&SE@SW 2023: 6th Workshop for Young Scientists in Computer Science & Software Engineering, February 2, 2024, Kryvyi Rih, Ukraine

vbazurin@gmail.com (V. M. Bazurin); o.pursky@knute.edu.ua (O. I. Pursky); karpenko26@i.ua (Y. M. Karpenko); t.pidhorna@knute.edu.ua (T. V. Pidhorna); a.nechepurenko@knute.edu.ua (A. I. Nechepourenko)
https://knute.edu.ua/blog/read/?pid=43549 (V. M. Bazurin); https://knute.edu.ua/blog/read/?pid=39582&uk (O. I. Pursky); https://zu.edu.ua/inyaz_kaf7.html (Y. M. Karpenko); https://knute.edu.ua/blog/read/?pid=44318&uk (T. V. Pidhorna); https://knute.edu.ua/blog/read/?pid=46356 (A. I. Nechepourenko)
0000-0002-6614-4889 (V. M. Bazurin); 0000-0002-1230-0305 (O. I. Pursky); 0000-0003-3630-5415 (Y. M. Karpenko); 0000-0002-1414-3489 (T. V. Pidhorna); 0009-0003-5828-0762 (A. I. Nechepourenko)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

CEUR-WS.org/Vol-3662/paper14.pdf

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

are advisable to use in the process of calculating the thermal insulation parameters of buildings and structures.

To calculate the parameters of thermal insulation, one can use general-purpose application software, automated design systems, or existing thermal conductivity calculators. The use of software of the specified categories has its advantages and disadvantages; in particular, they do not take into account the peculiarities of heat transfer in dielectric materials. In dielectrics, heat transfer processes occur due to phonon-phonon interaction without the participation of free electron diffusion [2, 3]. Therefore, it would be more appropriate to use a specialized software tool – a thermal resistance calculator for dielectric building materials.

The article aims to reveal the features of software implementation and the user interface of the thermal resistance calculator for dielectric building materials.

2. Computation method

One of the urgent and at the same time the most difficult problems of thermal physics is the task of researching energy transfer processes [4, 5, 6]. Data on thermal conductivity allow for obtaining important information for thermal physics. Dielectrics have several mechanisms that determine thermal conductivity, and a number of processes that limit the effectiveness of each of the mechanisms [7, 8, 9, 10]. The development of industrial technologies contributed to the creation of high-precision methods of thermal conductivity research, which made it possible to obtain more complete information about the nature of the forces acting in dielectric materials. However accurate quantitative prediction of thermal conductivity due to the complexity of describing specific mechanisms of heat transfer for real dielectrics is a difficult task. In the classical Debye model, the thermal conductivity Λ of dielectrics can be presented by the expression [2, 11]:

$$\Lambda = \frac{\hbar}{2\pi^2 v^2 k_b T^2} \int_a^{\omega_D} l(\omega)\omega^4 \frac{e^{\hbar\omega/k_B T}}{(e^{\hbar\omega/k_B T} - 1)^2} d\omega \tag{1}$$

where v is the sound velocity, ω_D is the Debye frequency, k_B is the Boltzmann's constant, $l(\omega)$ is the phonon mean free path, and $\hbar$ is the Planck's constant.

At the same time, today there are considerable deviations between the results of experimental studies of the dielectric materials thermal conductivity and classical theoretical predictions. At present, there is no generally accepted theoretical model that can effectively determine the contributions of different heat transfer mechanisms to the thermal conductivity of dielectrics [7, 8, 12, 13, 14]. Therefore, in practice, a simplified idea of the possibility of independent separation of the contributions of each of the phonon scattering mechanisms to the total thermal resistance and determination of the combined phonon relaxation time τ is often used [15, 16, 17]:

$$\tau^{-1} = \sum_i^n \tau_i^{-1} \tag{2}$$

where τ_i is the relaxation time for each of the phonon scattering mechanisms in the dielectric.

The possibility of using such a simplification in the inverse functions for kinetic coefficients was also theoretically substantiated in works [18, 19]. According to the results of these works,

if a quantity is a direct correlation function (in our case thermal conductivity) goes to infinity, then it is necessary to consider the behavior of its inverse quantity (inverse correlation function in our case is thermal resistance) and add to its thermodynamic value terms that describe spatiotemporal dispersion and additional corrections. If the inverse value goes to zero, then all correction terms are added to this zero value.

In the harmonic approximation [20], heat waves propagate in the crystal without interacting with each other. In this case, the arbitrary distribution of phonons does not become equilibrium even in the absence of a temperature gradient, since the mechanism of phonons relaxation to an equilibrium state is absent. The heat flow in this case exists indefinitely, and the thermal conductivity is infinite $\Lambda = \inf$, due to the absence of thermal resistance $W = 0$ ($\Lambda = 1/W$). But, as the experiment shows, the thermal conductivity of dielectrics is a finite value. Even in ideal crystals, the existence time of phonons is limited by the processes of phonon scattering, which arise as a result of the interaction potential anharmonicity of a real dielectric.

Thus, phonon anharmonicity is a necessary condition for finite thermal conductivity [2, 20]. In the framework of three-phonon interaction [2, 15], it is customary to distinguish *N*-processes and *U*-processes. *N*-processes during collisions do not change the value of the average quasi-momentum and by themselves do not lead to the occurrence of thermal resistance $W_N = 0$. After the *U*-process, there is a change in the direction of heat energy transfer, which ultimately leads to the emergence of thermal resistance $W_N = 0$ and repair of the phonons equilibrium distribution. In contrast to atomic crystals for dielectrics, in addition to phonon$-$phonon scattering, there are additional mechanisms of phonon scattering, in particular, boundary scattering of phonons, etc. [4, 7, 8, 15, 16]. That is, for dielectrics it is possible to write $\Lambda = 1/\sum W_i$ [3, 4] and assume that the contributions of different phonon scattering mechanisms to the total thermal resistance W are additive [8, 21]. Considering that different mechanisms of phonon scattering are present in dielectrics, the thermal resistance of the dielectric can be determined as:

$$W = \sum_{i=1}^{n} \frac{1}{W_i} \tag{3}$$

where W is the total thermal resistance, W_i are components of the total thermal resistance associated with different mechanisms of phonon scattering in dielectric materials [4, 21].

The formulas presented in works [22, 23, 24, 25, 26, 27] are used to calculate the values of thermal resistance for dielectric building materials. Calculations are performed under the assumption that the contributions of different phonon scattering mechanisms to the total thermal resistance of dielectric building materials are additive [2, 8].

For full thermal resistance W_R of thermally homogeneous and non-homogeneous opaque enclosing structures:

$$W_R = \frac{1}{\alpha_{in}} \sum_{i=1}^{n} \frac{\delta_i}{\lambda_i} \frac{1}{\alpha_{ext}} \tag{4}$$

where α_{in} is the heat exchange coefficient of the internal surface of the structure with the internal air; α_{ext} is the heat exchange coefficient of the outer surface of the structure with the outside air; δ_i is the thickness of the $i-$th material layer; λ_i is the coefficient of thermal conductivity of the $i-$th material layer.

For the thermal resistance W_{inc} of structures with defined values of the linear heat transfer coefficient of thermal-conducting inclusions:

$$W_{inc} = \frac{F_{tr} + \sum\limits_{i=1}^{n} F_i}{\frac{F_{tr}}{W_{tr}} + \sum\limits_{i=1}^{n} \frac{F_i}{W_i} + \sum\limits_{j=1}^{m} \lambda_j L_j} \tag{5}$$

where W_{tr} is the reduced heat transfer resistance of the transparent area which is taken depending on the characteristics of the double−glazed units (the distance between the layers of glass, the type of gas filling and the degree of blackness of the glass surface), F_{tr} is the area of the translucent part (m^2), W_i and F_i are thermal resistance and the area of the $i-$th opaque element, n is the number of opaque structural elements with certain values of W_i and F_i, k_j is the linear coefficient of thermal conductivity of the $j-$th inclusion, L_j is the linear size of the $j-$th opaque inclusion of the translucent structure, m is the number of opaque structural inclusion for which k_j must be determined.

In their article, Zahorodko et al. [28]. compare the performance of quantum-enhanced and classical machine learning algorithms in an experiment [28].

3. Results and discussion

Starting with the design of the thermal resistance calculator, we will analyze the functionality and interface of the most common programmes for this purpose. A search for relevant software found such software products as Rockwool U-value Calculator, Kingspan U-value Calculator and others. The Rockwool U-value Calculator is available on the website [29]. This software tool provides an opportunity to calculate the thermal resistance of walls, roof and floor.

Advantages of the Rockwool U-value Calculator:

- variability of structural and thermal insulation materials;
- intuitive interface;
- multilingual interface.

Disadvantages:

- non-compliance with the State Building Regulations of Ukraine (Rockwool U-value Calculator calculates thermal conductivity, but we need to calculate thermal resistance);
- uncontrollability (the owners of the resource can close access to this resource at any time, and its use in the educational process will become impossible).

The next software tool is the Kingspan U-value Calculator [30]. After registering the user on the website, he can choose the desired type of building and type of wall, determine the material of the internal equipment of the wall, the material and thickness of the main material, the type of mortar, the material of thermal insulation and its thickness, the thickness of the air layer, the material and thickness of the external decoration. This calculator has the same disadvantages as the previous one:

- thermal conductivity is calculated, not thermal resistance;
- lack of control over the resource on which the Kingspan U-value Calculator is located;
- a set of thermal insulation materials from Kingspan.

The analysis results show that it is impractical to use the mentioned thermal resistance calculators, since the method of their calculations does not correspond to the State Building Regulations [27]. That is, a significant drawback of these calculators is the use of another method of calculation.

A feature of thermal resistance calculations performed by civil engineers in Ukraine is that the calculation method and main coefficients are described in the State Building Regulations [27]. In these calculations, thermal resistance is calculated as the body's ability to prevent the spread of heat.

A feature of thermal resistance calculations performed by civil engineers in Ukraine is that the calculation method and main coefficients are described in the State Building Regulations [27]. In these calculations, thermal resistance is calculated as the body's ability to prevent the spread of heat, that is, as a process that is the inverse of thermal conductivity. Having analyzed the main calculations performed during the design of the thermal insulation coating of buildings and structures, it is possible to distinguish 4 main types of enclosing structures, and, accordingly, 4 main types of calculations [27]:

- calculation of the thermal resistance of the wall;
- calculation of the thermal resistance of the window;
- calculation of the thermal resistance of the attic floor;
- calculation of the thermal resistance of the roof overlap.

These calculations do not have fundamental differences, but use different limit values of thermal resistance coefficients and different structural and thermal insulation materials. Accordingly, the developed thermal resistance calculator must have appropriate functional modules. In addition, it is necessary to allocate an additional block of initialization of tabular data, which initializes the critical values of thermal resistances for various types of enclosing structures depending on the climatic zones of Ukraine. Let's consider these functional modules in more detail.

The *initial data initialization module* activates those critical values of thermal resistance that are characteristic of the relevant climatic zone of Ukraine. There are two such zones in Ukraine. In order to simplify this process, the user is given the opportunity to select the area in which the designed building will be constructed, and depending on the area, the program determines the climatic zone, and only after that activates the required values of thermal resistance for this zone. The specified thermal resistance values are output to the appropriate controls on the form. In the same module, basic classes are initialized and basic class instances are created.

The *module for calculating the wall thermal resistance* accepts the following input data (selected by the user using drop-down lists): material and thickness of the main wall, material, and thickness of thermal insulation, material, and thickness of internal plaster, material, and thickness of external cladding. In the handbooks, the coefficients of thermal conductivity of most building materials are given, but in this module, you can choose only those materials

that are used for the construction and thermal insulation of the wall. For example, there is no roofing material in the list of materials for walls (it is not used for waterproofing walls).

The *module for calculating the window thermal resistance* uses different input data, according to the materials used to make the windows. The tabular value of the thermal resistance of the window is different from that of the wall.

The *module for calculating the attic floor thermal resistance* simulates the corresponding calculations performed by the civil engineer. The user uses the drop-down lists to select the material of the main floor, the material of thermal insulation, and the material of the interior of the building, and with the help of sliders (TrackBar) sets the thickness of the corresponding structural elements. The programme calculates the thermal resistance and displays its value in the text field. Depending on whether the calculated value of thermal resistance corresponds to the permissible value for the given climatic zone, the text field is painted in green or red.

The *module for calculating the roof overlap thermal resistance* differs from the previous one only in a different set of materials (main support, thermal insulation, etc.), as well as in other values of heat exchange coefficients with the environment and the internal environment.

In the process of developing the thermal resistance calculator, an object-oriented approach was applied, according to which all information about a certain phenomenon of reality is concentrated in the object.

The class diagram of the “ThermoResist” programme is shown in figure 1.

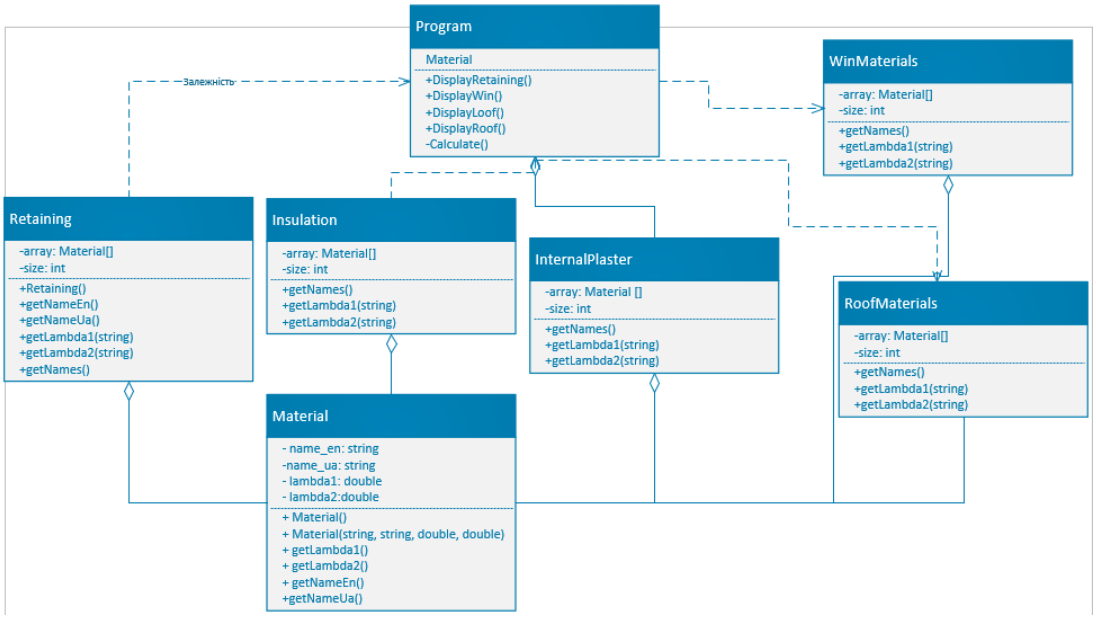


Figure 1: Class diagram of the “ThermoResist” thermal resistance calculator programme.

The programming language C♯ was chosen for the implementation of the programme. The creation and compilation of the software code was carried out in the Microsoft Visual Studio 2022 programme development environment. System requirements of the programme:

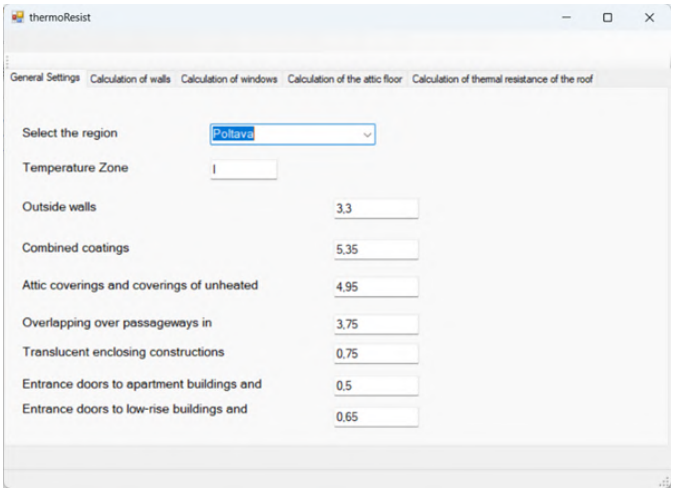


Figure 2: "ThermoResist" thermal resistance calculator interface.

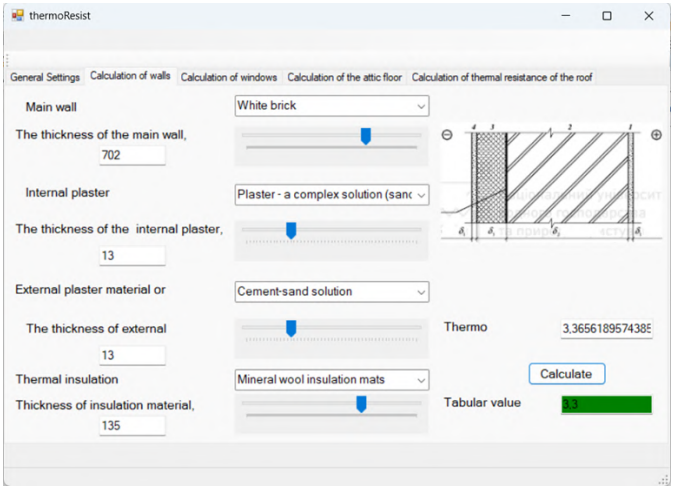


Figure 3: Calculation of the thermal resistance of the wall using the "ThermoResist" programme.

- monitor resolution 800x600;
- Microsoft Windows 10/11 operating system;
- .NET Framework 4.0 and higher.

The output of calculation results is accompanied by the coloring of the text field in green. This means that the thermal resistance of the wall has an acceptable value.

4. Conclusions

Digitization of society involves digitalization of all aspects of social life. The developed "ThermoResist" thermal resistance calculator for dielectric building structures has an intuitive in-

terface, a methodology, and a sequence of calculations that comply with the State Building Regulations in Ukraine. Thanks to this, basic computer user skills are enough to use this calculator, so it can be used by both students and civil engineers. Unlike many software products, it calculates thermal resistance according to the methodology given in the State Building Regulations and enables to calculation of the thermal resistance of the following categories of structures: walls, windows, attic floor, and roof according to indicators, specified in the State Building Regulations. The presented results will be useful for implementation in the educational process.

References

[1] S. Kraus, P. Jones, N. Kailer, A. Weinmann, N. Chaparro-Banegas, N. Roig-Tierno, Digital Transformation: An Overview of the Current State of the Art of Research, SAGE Open 11 (2021) 21582440211047576. doi:10.1177/21582440211047576.

[2] R. Berman, P. G. Klemens, Thermal conduction in solids, 1978.

[3] O. I. Purskii, N. N. Zholonko, V. A. Konstantinov, Heat transfer in the orientationally disordered phase of SF6, Low Temperature Physics 26 (2000) 278–281. doi:10.1063/1.593899.

[4] V. G. Manzhelii, Y. A. Freiman, Physics of cryocrystals, Amer Inst of Physics, 1997.

[5] C. Kittel, P. McEuen, Introduction to Solid State Physics, John Wiley & Sons, 2018.

[6] L. R. Ingersoll, An Introduction to the Mathematical Theory of Heat Conduction, Wentworth Press, 2019.

[7] V. A. Konstantinov, Heat transfer by low-frequency phonons and "diffusive" modes in molecular crystals, Low Temperature Physics 29 (2003) 422–428.

[8] O. I. Purskii, N. N. Zholonko, Heat transfer in the high-temperature phase of solid SF6, Physics of the Solid State 46 (2004) 2015–2020.

[9] C. Kittel, Quantum theory of solid states, Wiley, 2004.

[10] D. G. Cahill, S. K. Watson, R. O. Pohl, Lower limit to the thermal conductivity of disordered crystals, Physical Review B 46 (1992) 6131.

[11] C. Jeong, S. Datta, M. Lundstrom, Full dispersion versus Debye model evaluation of lattice thermal conductivity with a Landauer approach, Journal of Applied Physics 109 (2011) 073718. doi:10.1063/1.3567111.

[12] A. I. Krivchikov, O. O. Romantsova, O. A. Korolyuk, G. A. Vdovichenko, Y. V. Horbatenko, Specific features of heat transfer in the orientationally ordered phases of molecular crystals in the region with predominant phonon-phonon scattering, Low Temperature Physics 41 (2015) 551–556.

[13] A. J. H. McGaughey, M. Kaviany, Thermal conductivity decomposition and analysis using molecular dynamics simulations. Part I. Lennard-Jones argon, International Journal of Heat and Mass Transfer 47 (2004) 1783–1798.

[14] A. J. H. McGaughey, M. Kaviany, Thermal conductivity decomposition and analysis using molecular dynamics simulations: Part II. Complex silica structures, International Journal of Heat and Mass Transfer 47 (2004) 1799–1816.

[15] J. Callaway, Model for lattice thermal conductivity at low temperatures, Physical Review 113 (1959) 1046.
[16] Y. Zhang, First-principles Debye–Callaway approach to lattice thermal conductivity, Journal of Materiomics 2 (2016) 237–247.
[17] M. G. Holland, Analysis of lattice thermal conductivity, Physical review 132 (1963) 2461.
[18] D. N. Zubarev, Nonequilibrium Statistical Thermodynamics, Springer, New York, 1995.
[19] V. M. Sysoev, A. V. Chalyi, Correlation functions and dynamical structure factor of a nonisotropic medium near the critical point, Teoreticheskaya i Matematicheskaya Fizika 19 (1974) 283–288.
[20] L. P. Pitaevskii, E. M. Lifshitz, Physical Kinetics, volume 10, Butterworth–Heinemann, 2012.
[21] V. A. Konstantinov, V. P. Revyakin, V. V. Sagan, O. I. Pursky, V. M. Sysoev, Thermal conductivity of solid cyclohexane in orientationally ordered and disordered phases, Journal of Experimental and Theoretical Physics 112 (2011) 220–225.
[22] F. Akankali, H. Zefer Alibaba, Analysis on the thermal transmittance/conductance of building elements (roof, wall and floor) based on the different building materials used in Famagusta, Int. J. Civ. Struct. Eng. Res 6 (2019) 98–108.
[23] M. Belusko, F. Bruno, W. Saman, Investigation of the thermal resistance of timber attic spaces with reflective foil and bulk insulation, heat flow up, Applied energy 88 (2011) 127–137.
[24] B. M. Suleiman, Estimation of U-value of traditional North African houses, Applied Thermal Engineering 31 (2011) 1923–1928.
[25] X. Kong, J. S. McEwan, L. A. Bizo, T. M. Foster, An analysis of U-value as a measure of variability, The Psychological Record 67 (2017) 581–586.
[26] C. Underwood, F. Yik, Modelling methods for energy in buildings, John Wiley & Sons, 2008.
[27] H. Fareniuk, M. Timofeev, O. Oleksienko, SCN B.2.6-31:2021 Thermal insulation and energy efficiency of buildings, Minregion of Ukraine, 2022.
[28] P. V. Zahorodko, S. O. Semerikov, V. N. Soloviev, A. M. Striuk, M. I. Striuk, H. M. Shalatska, Comparisons of performance between quantum-enhanced and classical machine learning algorithms on the IBM Quantum Experience, in: Journal of Physics: Conference Series, volume 1840, IOP Publishing, 2021, p. 012021. doi:10.1088/1742-6596/1840/1/012021.
[29] Rockwool U-value Calculator, 2023. URL: https://bim.rockwool.co.uk/u_value/.
[30] U-value Calculations. Kingspan GB, 2023. URL: https://www.kingspan.com/gb/en/campaigns/u-value-calculations/.

Using the Three.js library to develop remote physical laboratory to investigate diffraction

Pavlo I. Chopyk[1], Vasyl P. Oleksiuk[1,2] and Oleksandr P. Chukhrai[1]

[1]Ternopil Volodymyr Hnatiuk National Pedagogical University, 2 M. Kryvonosa Str., Ternopil, Ukraine
[2]Institute for Digitalisation of Education of the NAES of Ukraine, 9 M. Berlynskoho Str., Kyiv, 04060, Ukraine

Abstract

The article examines the process of designing and developing a virtual (remote) physics laboratory. The authors have identified the requirements of the laboratory for studying the diffraction effect (visibility, accessibility, flexibility and visualization). The ability to set additional working parameters such as measurement error, material types and performance conditions. No additional software is required to carry out laboratory work, only an Internet connection and a browser.). The research analyzes the available digital tools for developing 3D models and ultimately selects the 3D library Theree.js. The researchers then proceed to create a suitable model for representing laboratory objects and processing experimental data. The virtual laboratory enables students to measure the line widths of the diffraction pattern when using monochromatic light (red, blue, green). It also enables the solution of the inverse problem, allowing determination and verification of the wavelength. Additionally, students may calculate measurement errors through the virtual laboratory.

Keywords

remote physical experiments, JavaScript, Three.js, remote lab

1. Introduction

Experiments and laboratory work are an essential part of modern physics education in secondary schools, colleges and universities. They facilitate learners' understanding of physical principles. Laboratory assignments aid students in gaining a better understanding of theoretical concepts by enabling them to observe the operation of physics principles in the natural world. Additionally, practical laboratory sessions allow teachers to assess students' knowledge and understanding. Lab exercises also facilitate student acquisition of skills in data manipulation, processing, and analysis. These skills are valuable not only for the study of physics but also for a wide range of other fields in science and technology. Laboratory experiments promote critical thinking and problem-solving and have the potential to stimulate an interest in science. Furthermore, practical skills are refined through lab work, such as operating scientific apparatus and practical abilities with potential applications in the real world [1].

CS&SE@SW 2023: 6th Workshop for Young Scientists in Computer Science & Software Engineering, February 2, 2024, Kryvyi Rih, Ukraine

chip.ukraine@gmail.com (P. I. Chopyk); oleksyuk@fizmat.tnpu.edu.ua (V. P. Oleksiuk); olexandrtera@gmail.com (O. P. Chukhrai)
https://tnpu.edu.ua/faculty/fizmat/chopik.php (P. I. Chopyk);
https://tnpu.edu.ua/faculty/fizmat/oleksyuk-vasil-petrovich.php (V. P. Oleksiuk)
0000-0001-6631-078X (P. I. Chopyk); 0000-0003-2206-8447 (V. P. Oleksiuk); 0009-0002-7075-3617 (O. P. Chukhrai)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

CEUR-WS.org/Vol-3662/paper23.pdf
CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

Currently, remote training is a common practice. Therefore, relevant software has been developed by many vendors and educators [2]. Although virtual laboratory work cannot always substitute traditional laboratory work. It has proven to be essential in certain circumstances [3, 4, 5]. The versatility offered by virtual labs allows students to work from any location, at any time [6]. Moreover, it eliminates the potential risks associated with handling dangerous or costly equipment, which could pose a threat to the safety and well-being of students. This facilitates the verification of results and the utilisation of modern technology. It is feasible to replicate experiments and substantiate findings within virtual laboratories. Researchers can experiment with concepts that are impractical or unfeasible within a conventional laboratory setting, such as exploring nuclear fission or the Big Bang theory.

The absence of practical experience and restricted engagement of students are drawbacks of virtual laboratories. Some concepts may have limited comprehension due to a lack of interaction with physical equipment. Furthermore, students may not adequately develop skills working with equipment. Therefore, it is advisable to incorporate a combined approach, combining both practical and virtual modelling. Virtual models aid students' readiness for laboratory experiments by facilitating comprehension and application of the material.

2. The model of a physical remote laboratory

To set up a remote physics laboratory, an educator can make use of readily available software or create their own. There are some free virtual laboratories such as [7, 8, 9, 10, 11, 12]

- PhET Interactive Simulations from the University of Colorado;
- Physics at School – HTML5-virtual environment ;
- Labster platform;
- CK-12 Interactive Physics for High School;
- Gizmos from Explore Learning.

All of them were created using HTML5, JavaScript, or Java technologies. Significant disadvantages that prevent the use of existing laboratories are the lack of monitoring of students' laboratory work, the inability to conduct measurements without displaying the finished results (in fact, simulations of physical processes are performed), and the perfection of the results, which does not allow for the assessment of measurement errors.

We decided to develop our service for running virtual physics labs. This study aims to systematise the design and development process of creating a remote physics laboratory for diffraction investigation.

Based on the analysis of studies by Holly et al. [13], Manabe et al. [14], Shepiliev et al. [15], Kiv et al. [16], Kanivets et al. [17] we have identified the following requirements for the service

- *Visibility.* The physical process should be fully visualised with the possibility of interactive control.
- *Accessibility.* No additional software is required to carry out laboratory work, only an Internet connection and a browser.
- *Flexibility.* The ability to set additional working parameters such as measurement error, material types and performance conditions.

- *Visualisation.* The appearance of virtual equipment should be similar to that of real laboratory equipment.

2.1. The frameworks, engines and libraries to developing physical remote laboratory

To establish remote laboratories, it is essential to carefully choose adequate development tools, which can include frameworks, engines, and 3D libraries. When considering the comparison of 3D graphics libraries, it is crucial to establish unambiguous standards to evaluate their aptitude and efficacy. The following are essential criteria to consider when evaluating 3D graphics libraries:

1. *Performance and rendering quality.* This criteria assesses the frame rate and rendering speed of intricate scenes, and the quality of graphics rendered, comprising lighting, shadows, and textures.
2. *Ease of use and learning curve.* The criteria are an assessment of the accessibility of tutorials, documentation, community support and ease of use of the library. It also includes the intuitiveness of the API and overall developer-friendliness.
3. *Flexibility and customization.* The assessment of this criterion involves determining the levels of ability to customize the rendering pipeline, effects and support for various 3D models, file formats, and asset management.
4. *Extensibility and modularity for integrating with other technologies.*
5. *Feature set and capabilities* are evaluated through the availability of essential 3D features and support for physics simulation, collisions, and particle systems. It also includes VR/AR support, if needed for the project.
6. *Community and ecosystem.* This criteria assesses the size, activity of the developer community, availability of third-party plugins, extensions, or add-ons, and frequency of updates and ongoing development.
7. *Cross-platform compatibility* provides for the definition the levels of support for various web browsers or platforms (desktop, mobile, VR devices), and the performance consistency across different devices and browsers.
8. *Integration and interoperability.* The criteria are an assessment of the compatibility with other popular frameworks, libraries, or technologies and ease of integration with existing workflows, tools, IDEs.
9. *Real-time collaboration and deployment* are evaluated through support for real-time collaboration among team members and smooth deployment and hosting options for web-based 3D applications.
10. *Project size and loading speed.* The criterion defines the impact of the library on the overall project size. It also includes booting time and efficiency in loading assets and scenes.
11. *License and open-source nature* is important in terms of material costs for project development. It involves assessing licensing terms (open-source, commercial, etc.), associated costs, and availability of a free or trial version for evaluation.
12. *Use cases and education adoption* This criteria assesses the relevance of the library to physics education tasks and examples of successful projects or applications built using

the library. The criterion is interdisciplinary. To assess it, there is a demand for training STEM education specialists who are competent in physics and able to develop and deploy virtual and augmented reality software [18].

There are some well-known libraries and frameworks that can be used to create 3D graphics and animations. These include Three.js, Babylon.js, A-Frame and Cannon.js. Based on the literature reviews by [19, 14, 20] and discussions with practising developers we have conducted a comparative analysis of these 3D graphics libraries. A 5-point scale was used to evaluate the criteria (5 points – maximum compliance of the framework with the criterion, 1 point – minimum compliance of the framework with the criterion). The results of analysis are given in table 1.

Table 1
Compare the capabilities of some libraries and frameworks that can be used to create 3D graphics and animation.

Criteria	Framework			
	Three.js	Babylon.js	A-Frame	Cannon.js
performance and rendering quality	4	5	3	2
ease of use and learning curve	4	4	5	4
flexibility and customization	5	4	3	2
feature set and capabilities	5	5	3	2
community and ecosystem	5	5	4	3
cross-platform compatibility	5	4	4	4
integration and interoperability	4	3	3	4
real-time collaboration and deployment	4	4	3	1
project size and loading speed	4	3	4	4
license and open-source nature	5	5	5	5
use cases and education adoption	4	3	5	3

So, we chose the Three.js library to develop a virtual physical laboratory to investigate the diffraction effect.

The process of designing our virtual laboratory to investigate diffraction predicted took place during such stages.

1. Formulation of the physical problem.
2. Selecting development tools and obtaining the required libraries.
3. The creation and design of a laboratory model.
4. Implementation and testing.

The diffraction grating's constant, d, the distance from the grating to the screen, L, and the diffraction order, m, are provided. Students are to measure the distance from the centre to the maximum l. From these two distances, they have to determine the angle β. Since the β angle is small, $tg(\beta) \approx sin(\beta) \approx \beta$ (figure 1.) We assume that a plane light wave is normally incident on the grating.

There are several steps in the process of conducting an experiment, whether in reality or virtually.

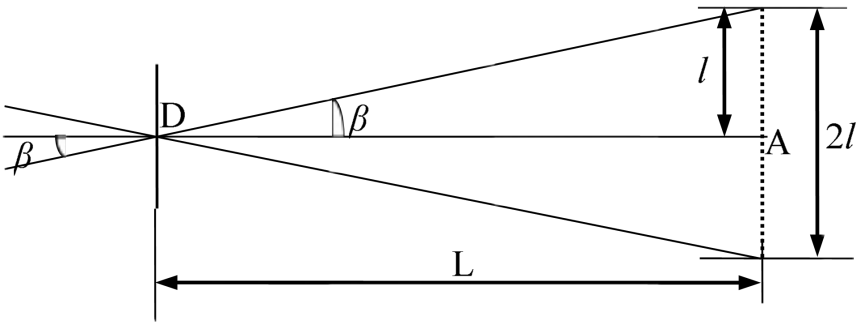


Figure 1: The physics problem of diffraction.

1. The diffraction spectrum on the screen should be captured accurately. Finally, the spectrum should be analysed.
2. Position the diffraction grating at a specific distance, L, from the screen.
3. Calculate the distance l, which represents the distance between the center of the line in the given order's spectrum and the zero spectrum, as illustrated in figure 1.
4. Employ formula (1) to calculate the wavelength λ.

$$\lambda = \frac{d \cdot l}{n \cdot L} \tag{1}$$

   where d is the distance between the adjacent slits (grating period), n is the diffraction order (an integer).
5. Alter the distance between the screen and the grating repeat step 3.

2.2. Discussing the design and development of a remote laboratory

As mentioned above, we chose the Three.js library. Additionally, we employed Bootstrap, an open-source CSS framework designed for front-end web development. Therefore we have obtained the necessary libraries, such as three.js-master and bootstrap-dist and uploaded them to the web-server folder. The virtual laboratory model is shown in figure 2.

To actually be able to display diffraction with three.js, we need to create scene, camera and renderer. All objects in our virtual lab are placed in a `<div>` container on the canvas. The left menu contains the parameters of the laboratory objects such as switching on the lamp, activating light filters, moving the dynamic screen, changing the camera position, diffraction grating step, moving the pointer position (see figure 3).

There are a few different cameras in three.js. We used a PerspectiveCamera. It is designed to mimic the way the human eye sees [21]. Based on this mode, 3 camera instances have been created. They allow the student to see the phenomenon in three modes

1. *General view.* It shows a view of the equipment for studying diffraction in three-dimensional space.

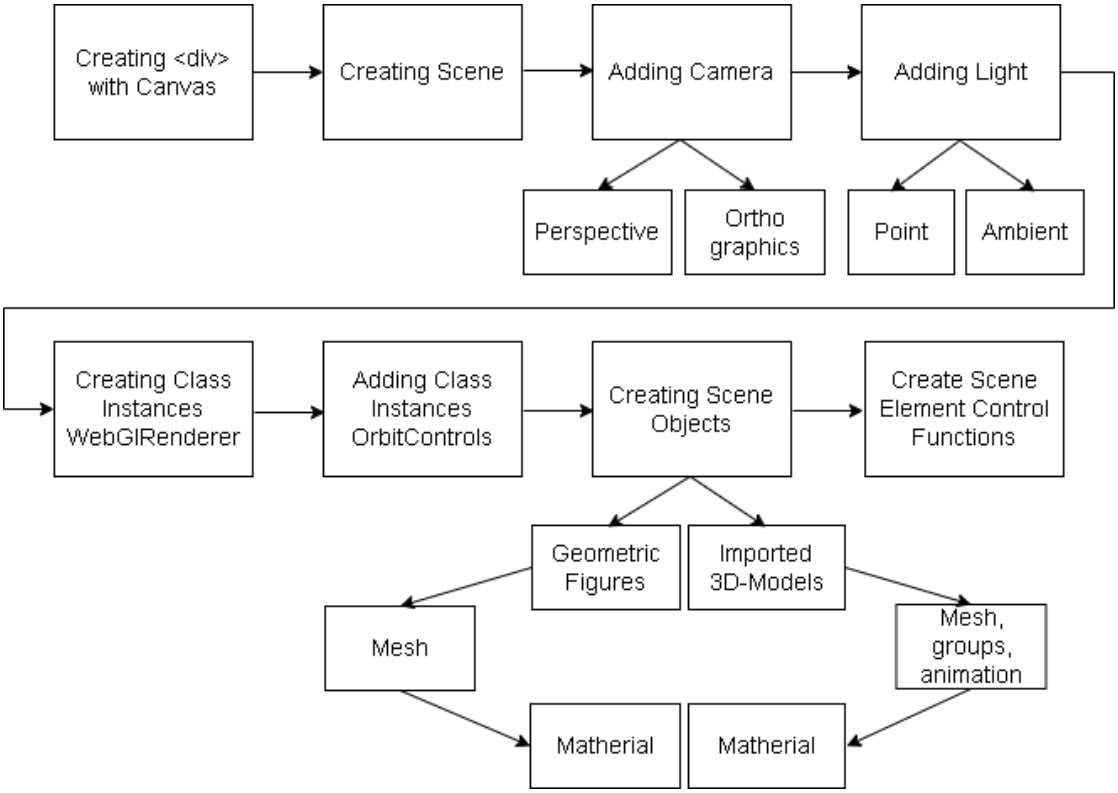


Figure 2: The model of virtual laboratory to investigate diffraction

2. *Observation mode.* It shows the diffraction pattern. This mode allows you to set the pointer.
3. Distance measurement mode. It provides rulers for measuring distance l from the formula 1.

The code 1 was used for the creation of a camera.

Listing 1: Camera creation example

```
camera=new THREE.PerspectiveCamera(45, webgl_width/webgl_height,
                                   1, 10000);
camera.position.set(cameraPos1.px,cameraPos1.py,cameraPos1.pz);
camera.rotation.set(cameraPos1.rx,cameraPos1.ry,cameraPos1.rz);
camera.name = "Camera1";
scene.add(camera);
```

To create the lighting, we used 3D objects such as PointLight and AmbientLight. The first one is emitted from a single point in all directions. The second globally illuminates all objects in the scene equally [21]. PointLight is used for the lamp, and AmbientLight is the background lighting for the entire scene.

Then an instance of the WebGLRenderer class was created with parameters such as

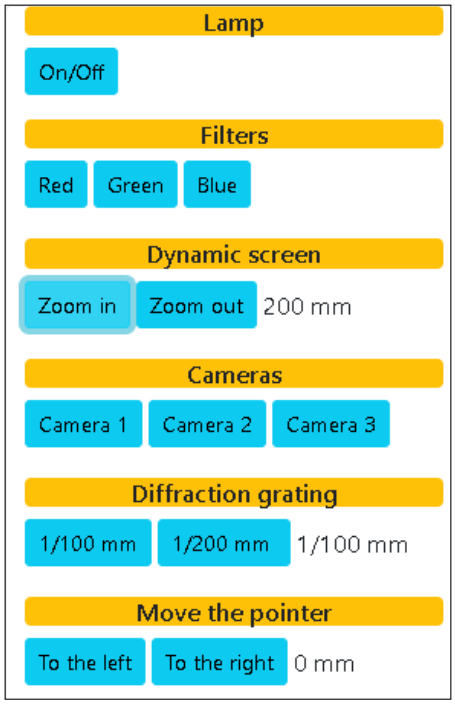


Figure 3: The menu of virtual laboratory to investigate diffraction.

- alpha controls the default opacity value. Its value was set to zero;
- antialias is used to reset the internal WebGL state;
- domElement is a canvas where the renderer draws its output;
- setSize sets dimension of the scene;

The figure 3 displays the menu for the virtual laboratory designed for investigating diffraction.

In order to enable camera movement around a grating, an instance of the OrbitControls class was instantiated. The properties of this instance determine the maximum and minimum distance of the dynamic screen and the angle‘s limit of vertical rotation (see code 2).

Listing 2: An instance of the OrbitControls class creation example

```
controls = new OrbitControls(camera, renderer.domElement);
controls.minDistance = 2;
controls.maxDistance = 25;
controls.maxPolarAngle = Math.PI / 2;
controls.target.set(0, 4, 0);
```

The scene objects were created by using geometric shapes and importing 3D models. BoxGeometry was used to create the main objects in the scene. These include the table, ruler, dynamic screen, stand and diffraction grating. It is a geometry class for a rectangular box with a given width, height and depth. The view of these objects from camera 1 is shown in figure 4.

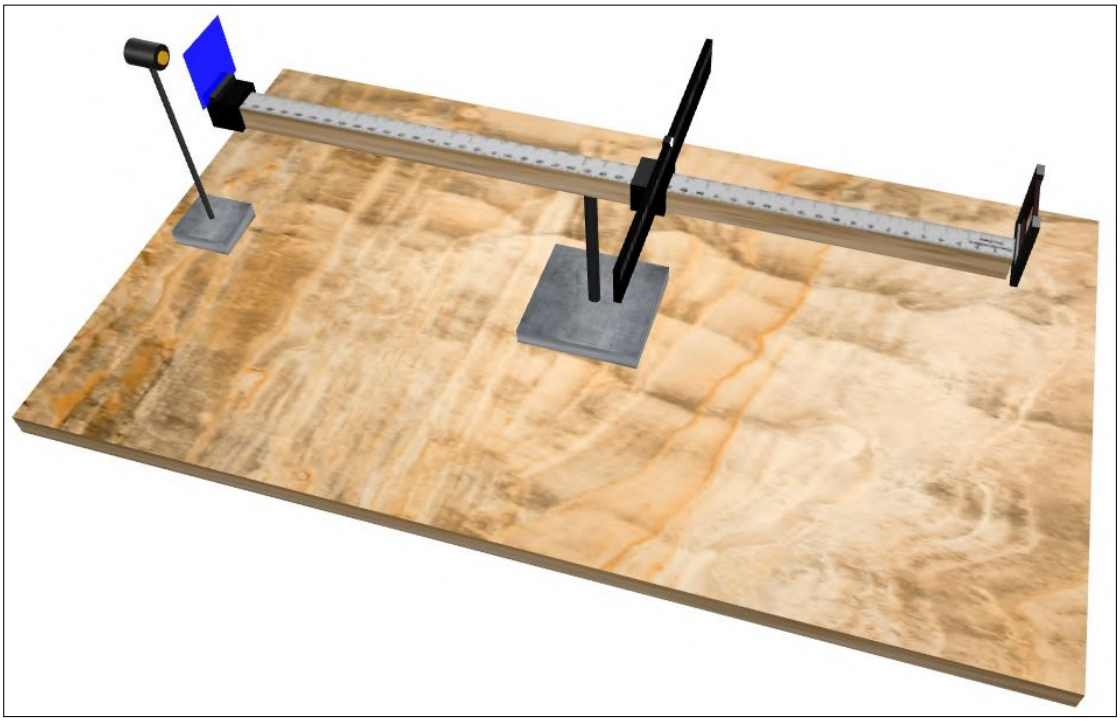


Figure 4: The view of the virtual laboratory from camera #1.

The materials from which these objects are made were created using the Mesh class. It represents triangular polygon mesh-based objects [21]. The instances of the MeshStandardMaterial class and its extensions MeshPhysicalMaterial and MeshPhongMaterial were used for creating some materials such as metal (black and silver), glass, plastic and table. For example, a diffraction grating in three dimensions was created using the code in listing 3.

Listing 3: The code for creating diffraction grating

```
var GratingGlassBox = new THREE.Mesh(new THREE.PlaneGeometry(0.7,
                                                    0.6), materialGlass);
GratingGlassBox.name = 'GratingGlassBox';
var GratingLftBox = new THREE.Mesh(geometryGratingLftRghtBox,
                                                    materialGratingPlastic);
var GratingRghtBox = new THREE.Mesh(geometryGratingLftRghtBox,
                                                    materialGratingPlastic);
var GratingUpBox = new THREE.Mesh(geometryGratingUpDwnBox,
                                                    materialGratingPlastic);
var GratingDwnBox = new THREE.Mesh(geometryGratingUpDwnBox,
                                                    materialGratingPlastic);
```

Other laboratory items were produced similarly, including a colour filter, mobile pointer, and dynamic screen ruler. They allow students to change the diffraction effect and measure the

distance between the centre of the spectral line of a given order and the spectral line at zero order. To utilize the dynamic screen ruler, employ camera view #3 figure 5.

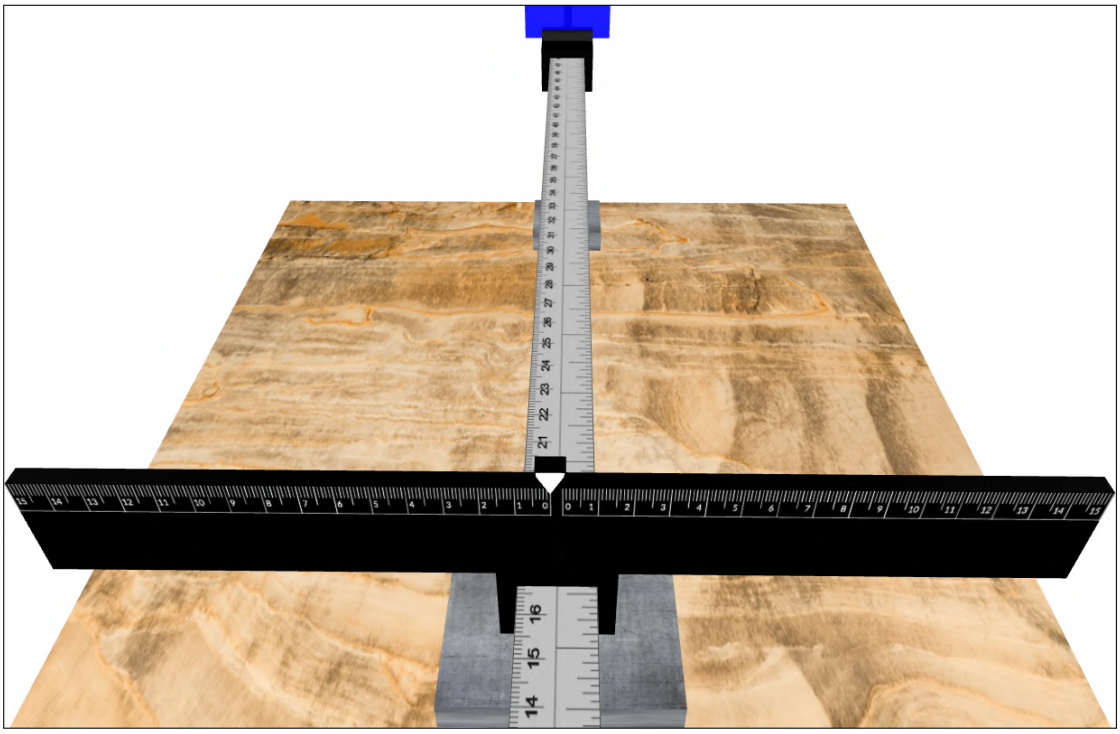


Figure 5: The view of the virtual laboratory from camera #3.

We used three groups to create complex objects such as a lamp and a tripod. They combined graphic primitives and imported images. The code for forming the lamp can be found in the listing 4.

Listing 4: The code for forming the lamp

```
{LampBase.position.set(0, 0, 0);
    LampVerticalCylinder1.position.set(0, 1.8, 0);
    LampHorizontalCylinder1.position.set(0, 3.58, 0);
    LampHorizontalCylinder1.rotation.x = 1.6;
    LampHorizontalCylinder1.rotation.x = 1.6;
    LampHorizontalColorCylinder.position.set(0, 3.572, 0.275);
    LampHorizontalColorCylinder.rotation.x = 1.6;
    GroupLamp.add(LampBase);
    GroupLamp.add(LampVerticalCylinder1);
    GroupLamp.add(LampHorizontalCylinder1);
    GroupLamp.add(LampHorizontalColorCylinder);
    GroupLamp.position.set(0, 1.15, -7);
    }
```

The colour filter was created as a Mesh object with the values THREE.BoxGeometry and THREE.MeshStandardMaterial (see listing 5).

Listing 5: The code for forming blue filter

```
var FilterGlass = new THREE.Mesh(
     new THREE.BoxGeometry(1, 1, 0.01),
     new THREE.MeshStandardMaterial({color: 0x0000ff,
                                     emissive: 0x0000ff,
                                     transparent: true,
                                     opacity: 0.9})
```

The calcPosSpectr() method computes the distance between the centre of the spectral line of a given order and the spectral line at zero order for the three colours (red, green, and blue) utilizing the formula (1). For example, the calculation for the first green line from the left is as follows (see listing 6)

Listing 6: The fragment of method for calculation the distance between the spectral lines

```
function calcPosSpectr()
{
    var elemLine = GroupDinamLine.getObjectByName
                    ('DinamLineLeftBox');
    var width_dinam_line = (getWidth(elemLine)/150);
    var PR = ((currentPosDinamBox/1000) /
                currentGrattingSize) * 1000;
    var gl1 = spectr.rw * 1 * PR;
 }
```

The first two lines of the calcPosSpectr() method convert matrix distances into screen coordinates. The constant 150 is the length of the ruler on the dynamic screen. In code (5), the variable PR is equal to $\frac{L}{d}$, spectr.gw is the width of the first green line.

Students are expected to perform their own calculations, including using cloud-based platforms such as CoCalc, Maxima, etc [22, 23, 24].

As a result, the diffraction effect can be observed in camera #2 (see figure 6). It also allows the student to measure the position and width of the spectral lines. In addition, boundary conditions are imposed on the positions of the ruler and pointer to ensure that they do not exceed the permitted limits. The displacement of the ruler is exactly 1mm. This was achieved by editing the texture of the ruler itself.

In the virtual lab, we made sure that the measurement result was not artificially tied to a particular value of physical quantities. This is why the diffraction pattern of the slit on the screen is slightly blurred. This makes it possible to obtain different values of distance during visual observation and to calculate the measurement error accordingly.

The real diffraction spectrum can be seen in figure 7. As you can see, the image is rather unclear and it is difficult to determine the position of the spectrum by looking through the diffraction grating. For this purpose, a pointer is provided in the virtual work and it is possible to switch to the view of camera #3 to measure distances more accurately.

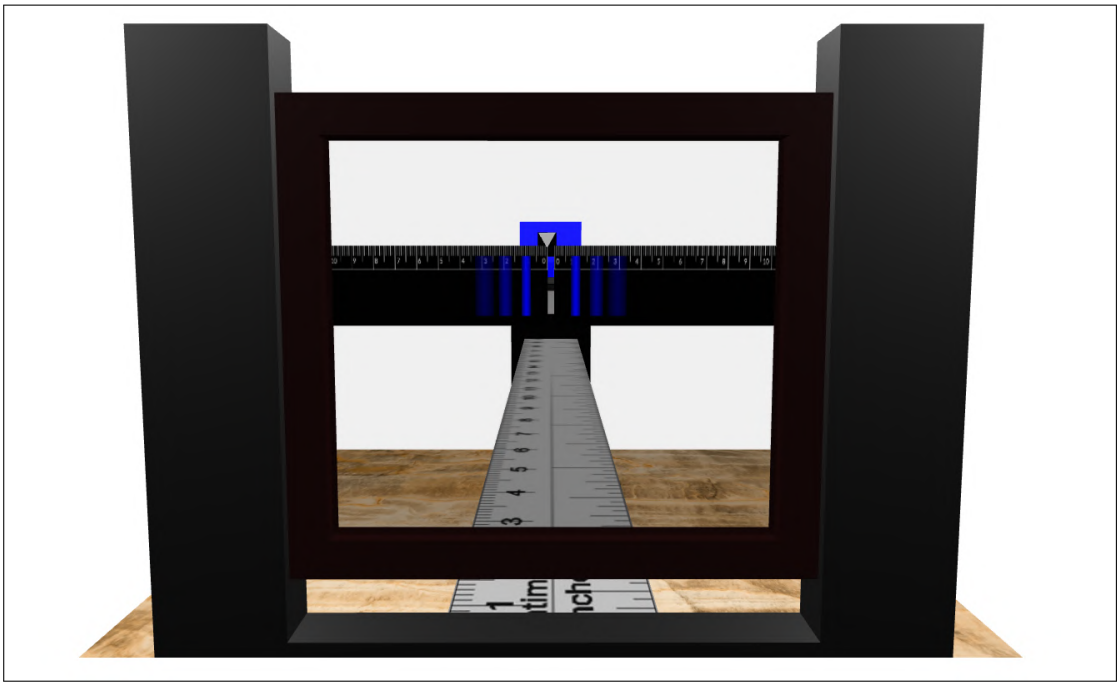


Figure 6: The view of the virtual laboratory from camera #2.

Figure 7: Real diffraction pattern.

We compare the results of the virtual and real experiments. When students used a blue light filter during virtual laboratory work, the following data were obtained: light wavelength $\lambda = 485 \pm 12$ nm with a relative measurement error of 2.3 %. During work on real equipment, the light wavelength was $\lambda = 472 \pm 25$ nm with a relative measurement error of 5.3 %. As we can see, the results within the margin of error coincide and correspond to the wavelength of blue light.

3. Conclusions

Developing virtual physics labs is a real problem. Modern digital tools such as frameworks and programming language libraries make this possible. Based on comparative analysis was selected one of these tools is Three.js. It is a JavaScript 3D library. In this research, the design and development of a virtual physics laboratory for studying the effects of diffraction using Three.js was carried out. The lab development process took place in several stages, such as creating scene, adding camera, creating main instances of classes (WebGLRenderer, OrbitControls), creating scene objects, and defining methods for calculating distances between spectral lines. The developed lab can be useful for blended learning in physics at colleges and universities. It can also be combined with real installations for the study of optics. The disadvantage of this virtual laboratory is the requirement of the Three.js library for modern hardware of the student's or teacher's device. This can be a limitation when working with older computers.

Three laboratory virtual works out of five, which are provided by the curriculum of the "Physics" course covering different sections (mechanics, molecular physics and thermodynamics), were developed using this method. In the future, it is planned to complete the development of the remaining works and combine them into a single integrated environment. This will allow students, in addition to performing work, to collect and process experimental data and prepare reports. We see prospects for further research, such as development of a version of this laboratory to provide student collaboration and its integration with modern learning management systems.

References

[1] E. K. Faulconer, A. B. Gruss, A review to weigh the pros and cons of online, remote, and distance science laboratory experiences, International Review of Research in Open and Distance Learning 19 (2018) 155–168. doi:10.19173/irrodl.v19i2.3386.

[2] Laboratory Experiences and Student Learning, in: S. R. Singer, M. L. Hilton, H. A. Schweingruber (Eds.), America's Lab Report: Investigations in High School Science, Washington, DC, 2006, pp. 75–115. doi:10.17226/11311.

[3] A. Striuk, O. Rybalchenko, S. Bilashenko, Development and using of a virtual laboratory to study the graph algorithms for bachelors of software engineering, CEUR Workshop Proceedings 2732 (2020) 974–983. URL: https://ceur-ws.org/Vol-2732/20200974.pdf.

[4] V. Oleksiuk, O. Spirin, The Experience of Using Cloud Labs in Teaching Linux Operating System, in: O. Ignatenko, V. Kharchenko, V. Kobets, H. Kravtsov, Y. Tarasich, V. Ermolayev,

D. Esteban, V. Yakovyna, A. Spivakovsky (Eds.), ICTERI 2021 Workshops, Springer International Publishing, Cham, 2022, pp. 281–291. doi:10.1007/978-3-031-14841-5_18.
[5] N. Balyk, V. Oleksiuk, Y. Vasylenko, G. Shmyger, Designing of Virtual Cloud Labs for the Learning Cisco CyberSecurity Operations Course, CEUR Workshop Proceedings 2393 (2019) 960–967. URL: https://ceur-ws.org/Vol-2393/paper_338.pdf.
[6] P. P. Nechypurenko, S. O. Semerikov, O. Y. Pokhliestova, An augmented reality-based virtual chemistry laboratory to support educational and research activities of 11th grade students, Educational Dimension 8 (2023) 240–264. doi:10.31812/educdim.4446.
[7] C. Chotimah, Festiyed, A meta-analysis of the effects of using PhET interactive simulations on student's worksheets toward senior high school students learning result of physics, Journal of Physics: Conference Series 1481 (2020) 012093. doi:10.1088/1742-6596/1481/1/012093.
[8] S. Siswoyo, D. Muliyati, Teaching high school physics using PhET interactive simulation, AIP Conference Proceedings 2331 (2021) 030003. doi:10.1063/5.0041657.
[9] V. Vascak, Physics at school, 2023. URL: https://vascak.cz/.
[10] S. K. Damekova, N. N. Shuyushbayeva, N. K. Tanasheva, G. S. Ahayeva, A. K. Kaliyeva, Additional education of schoolchildren in physics at the Children's University, Bulletin of the University of Karaganda-Physics 1 (2021) 63–70. doi:10.31489/2021Ph1/63-70.
[11] CK-12, CK-12 Interactive Physics for High School, 2023. URL: https://flexbooks.ck12.org/cbook/ck-12-physics-flexbook-2.0/.
[12] ExploreLearning, Online simulations for physics, 2023. URL: https://gizmos.explorelearning.com/physics-collection.
[13] M. Holly, J. Pirker, S. Resch, S. Brettschuh, C. Gütl, Designing VR Experiences – Expectations for Teaching and Learning in VR, Educational Technology and Society 24 (2021) 107–119. URL: https://www.jstor.org/stable/27004935.
[14] S. Manabe, C. Mouri, H. Tenzou, T. Kasuga, K. Motoki, Development of an AR chart tool for understanding invisible physical phenomena, AIP Conference Proceedings 1733 (2016) 020001. doi:10.1063/1.4948819.
[15] D. S. Shepiliev, S. O. Semerikov, Y. V. Yechkalo, V. V. Tkachuk, O. M. Markova, Y. O. Modlo, I. S. Mintii, M. M. Mintii, T. V. Selivanova, N. K. Maksyshko, T. A. Vakaliuk, V. V. Osadchyi, R. O. Tarasenko, S. M. Amelina, A. E. Kiv, Development of career guidance quests using WebAR, Journal of Physics: Conference Series 1840 (2021) 012028. doi:10.1088/1742-6596/1840/1/012028.
[16] A. E. Kiv, V. V. Bilous, D. M. Bodnenko, D. V. Horbatovskyi, O. S. Lytvyn, V. V. Proshkin, The development and use of mobile app AR Physics in physics teaching at the university, CEUR Workshop Proceedings 2898 (2021) 197–212. URL: https://ceur-ws.org/Vol-2898/paper11.pdf.
[17] O. V. Kanivets, I. M. Kanivets, T. M. Gorda, O. V. Gorbenko, A. O. Kelemesh, Using a mobile application to teach students to measure with a micrometer during remote laboratory work, CEUR Workshop Proceedings 3364 (2023) 87–107. URL: https://ceur-ws.org/Vol-3364/paper08.pdf.
[18] S. O. Semerikov, M. M. Mintii, I. S. Mintii, Review of the course "Development of Virtual and Augmented Reality Software" for STEM teachers: Implementation results and improvement potentials, CEUR Workshop Proceedings 2898 (2021) 159–177. URL: https://ceur-ws.org/

Vol-2898/paper09.pdf.
[19] D. S. Shepiliev, Y. O. Modlo, Y. V. Yechkalo, V. V. Tkachuk, M. M. Mintii, I. S. Mintii, O. M. Markova, T. V. Selivanova, O. M. Drashko, O. O. Kalinichenko, T. A. Vakaliuk, V. V. Osadchyi, S. O. Semerikov, WebAR development tools: An overview, CEUR Workshop Proceedings 2832 (2020) 84–93. URL: https://ceur-ws.org/Vol-2832/paper12.pdf.
[20] T. A. Vakaliuk, S. I. Pochtoviuk, Analysis of tools for the development of augmented reality technologies, in: Proceedings of the 4th International Workshop on Augmented Reality in Education (AREdu 2021), volume 2898, 2021, pp. 119–130. URL: https://ceur-ws.org/Vol-2898/paper06.pdf.
[21] R. Cabello, Three.js docs, 2023. URL: https://threejs.org/docs/.
[22] P. V. Merzlykin, M. V. Marienko, S. V. Shokaliuk, CoCalc: an integrated environment for open science education in informatics and mathematics, CEUR Workshop Proceedings 3482 (2023) 39–53. URL: https://ceur-ws.org/Vol-3482/paper025.pdf.
[23] S. Papadakis, A. E. Kiv, H. M. Kravtsov, V. V. Osadchyi, M. V. Marienko, O. P. Pinchuk, M. P. Shyshkina, O. M. Sokolyuk, I. S. Mintii, T. A. Vakaliuk, A. M. Striuk, S. O. Semerikov, Revolutionizing education: using computer simulation and cloud-based smart technology to facilitate successful open learning, CEUR Workshop Proceedings 3358 (2023) 1–18. URL: https://ceur-ws.org/Vol-3358/paper00.pdf.
[24] M. Shyshkina, U. Kohut, M. Popel, The comparative analysis of the cloud-based learning components delivering access to mathematical software, CEUR Workshop Proceedings 2393 (2019) 93 – 104. URL: https://ceur-ws.org/Vol-2393/paper_241.pdf.