



ICT-2023

28 листопада - 01 грудня
Харків 2023

12^a Міжнародна науково-технічна конференція
«Інформаційні системи та технології»
(ICT-2023)

Матеріали конференції

Частина 2. Молодіжна секція.



Харків 2023

Міністерство освіти та науки України
Національна академія наук України
Люблінський відділ Польської Академії Наук
Представництво „Польська академія наук” у Києві
Харківський національний університет радіоелектроніки
AGH науково-технологічний університет ім. Ст. Сташіца в Кракові
Прикарпатський національний університет ім. В. Стефаника
Національний університет кораблебудування ім. адмірала Макарова
Одеський державний екологічний університет
Національний університет Запорізька політехніка
Академія Наук Прикладної Радіоелектроніки
Українська нафтогазова академія
Українська Федерація Інформатики

Інформаційні системи та технології ІСТ-2023

Матеріали 12-ї Міжнародної науково-технічної конференції

Частина 2. Молодіжна секція

**28 листопада 2023 – 01 грудня 2023 р.
Харків, Україна**

Харків 2023

Ministry of education and science of Ukraine
The National Academy of Sciences of Ukraine
Polish academy of science, branch in Lublin
Representative office „Polish Academy of Science” in Kyiv
Kharkiv National University of Radio Electronics
Science and Technological University Academy of Mining and Metallurgy
after St. Staszic in Krakow
Stefanik Precarpathian National University
Admiral Makarov National University of Shipbuilding
Odesa State Environmental University
National University "Zaporizhzhia Polytechnic"
Academy of Applied Radioelectronic Sciences
Ukrainian Oil and Gas Academy
Ukrainian Federation of Informatics

Information Systems and Technologies IST-2023

**Proceedings
of the 12-th International Scientific and Technical Conference**

Part 2. Youth section

**November 28, 2023 – December 01, 2023
Kharkiv, Ukraine**

Kharkiv 2023

УДК: 004.9

Наукові редактори: *В.В. Безкоровайний - доктор технічних наук, професор (Харківський національний університет радіоелектроніки)*
L. Petryshyn - доктор технічних наук, професор (AGH University of Science and Technology; Прикарпатський національний університет ім. В. Стефаника);
З.В. Дудар – кандидат технічних наук, професор (Харківський національний університет радіоелектроніки)
Ю.В. Міщеряков.– кандидат технічних наук, доцент (Харківський національний університет радіоелектроніки)

Матеріали опубліковано в авторській редакції.

Інформаційні системи та технології: матеріали 12-ї Міжнародної науково-технічної конференції. Частина 2. Молодіжна секція, Харків, 28 листопада 2023 – 01 грудня 2023 року / наук. ред. В.В. Безкоровайний, L. Petryshyn, З.В. Дудар, Ю.В. Міщеряков. – Х.: ХНУРЕ, 2023. – 80 с.

Збірка містить матеріали Міжнародної науково-технічної конференції з проблем сучасних інформаційних систем та технологій.

Роботи представляють інтерес для фахівців, науковців, аспірантів, студентів, діяльність яких пов'язана з розробкою та впровадженням сучасних інформаційних систем і технологій.

© Харківський національний університет
радіоелектроніки, 2023

© Автори матеріалів, 2023

Application of clonal selection principle in neural architecture search

Dmytro Darahan^a, Oleksandr Sahaidachnyi^a, Nataliia Serdiuk^a

^aKharkiv National University of Radioelectronics, Nauky ave., 14, Kharkiv, 61166, Ukraine

Abstract

This paper describes the application of artificial immune systems in the domain of automatic machine learning. A method of applying the clonal selection algorithm for solving the task of searching for neural architecture is proposed. The prospects of using the described approach, its advantages and disadvantages are analyzed, and the future research subjects are outlined.

Keywords ¹

AutoML, artificial intelligence, artificial immune systems, deep neural networks, neural architecture search

1. Introduction

Automated machine learning, also known as AutoML, is the process of automating the time-consuming tasks of developing machine learning models. It allows to creation of machine learning models with high scalability, efficiency, and productivity, preserving model quality.

Depending on the essence of the problem, a deep learning expert may need tens to hundreds of experiments to find the balance between neural network architecture and hyperparameters. Such searches, as a rule, require a large amount of computing resources.

AutoML potentially includes all steps from starting with a raw dataset to building a deployment-ready machine learning model. AutoML was proposed as an AI-based solution to the growing problem of machine learning applications. Common techniques used in AutoML include hyperparameter optimization, meta-learning, and neural architecture discovery.

A subfield of AutoML, Neural Architecture Search (NAS) is a technique for finding the best neural network architecture for a given problem. NAS automates the design of deep neural networks, providing higher performance and less loss than manual architecture search.

The general idea of NAS is to choose the optimal architecture from all possible solutions. The selection algorithm is based on the search strategy, which, in turn, depends on the objective evaluation scheme.

There are actual problems of the application of artificial immune networks for tasks of automated machine learning. AIS uses ideas gleaned from immunology to develop adaptive systems capable of forming a wide range of tasks in various areas of scientific activity. One such task can be the neural architecture search.

2. Proposal of conception

Neural architecture search consists of three main components, which can be classified in terms of the type and complexity of algorithms, on which basis they are implemented: search space, search strategy, and evaluation strategy.

¹Proceedings of the *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine

EMAIL: dmytro.darahan@nure.ua (D. Darahan); oleksandr.sahaidachnyi@nure.ua (O. Sahaidachnyi);

nataliya.serdyuk@nure.ua (N. Serdiuk)

ORCID: 0000-0003-4918-4579(D. Darahan); 0000-0002-0107-4365 (N. Serdiuk)



The NAS search space defines what type of architecture can be discovered. It is defined by a set of operations that define the overall structure of the network, the type of units or blocks that define the layers, and the allowable communication between layers to create architectures.

The search strategy determines how the NAS algorithm experiments with different architectures. Based on a sample of the network population, the algorithm optimizes the model's performance metrics as a reward for generating a high-performance architecture output. Types of search algorithms include random search, evolutionary methods, Bayesian approaches, and reinforcement learning.

During NAS search, the algorithm learns, evaluates, tests, and compares performance before selecting the optimal neural network. Several strategies can be used to reduce the cost of evaluating deep learning models, including proxy tasks, low precision performance evaluation, learning curve extrapolation, network morphism

The clonal selection principle explains how the adaptive immune system deals with external foreign antibodies. The main elements, antibodies, and antigens are points in the L-dimensional space of forms, in which the measurements correspond to a set of features essential for the recognition of antigens by antibodies.

Antibodies recognize foreign antigens with a certain affinity. Affinity is a characteristic that quantitatively describes the strength of the interaction between an antigen and an antibody. Antibodies with high affinity are selected for cloning and multiple antibody clones are produced. Cloning – obtaining a given number of identical copies of the original antibody. Antibody clones undergo a mutation process. The mutation process leads to the fact that the newly created antibodies have a better affinity for a given antigen than the original selected antibody. Antibodies with improved affinity are stored in memory. Newly formed antibodies are removed if they have low affinity.

We propose to use the clonal selection algorithm as a search strategy for the NAS.

As an antigen, should be used the sought hyperparameters of the neural network in the Hamming space of forms. As a measure of affinity and an evaluation strategy, the neural network training metric will act at the testing stage, which must be optimized at each generation of the artificial immune system. To save computing resources on low-affinity antibodies, we suggest using early stopping and training coefficient planning during neural network training. Also, to speed up the search for the optimal network, it is planned to use proportional mutation and proportional cloning. The stop criterion for for the immune network is the lack of improvement in affinity above a certain threshold value during a certain, small number of generations.

The main problem of this approach is common to all NAS implementations - high computing costs, the need to correctly define the search space - the number and type of levels, their connections, and other hyperparameters of the network, etc.

3. Conclusion

We have proposed a conception of the application of the clonal selection algorithm for tasks of automatic search of neural architecture. We have considered the potential advantages and disadvantages of this approach. Our proposal should preserve all the advantages of NAS, i.e. accelerate and optimize the development of neural networks. At the same time, our approach remains prone to high costs of computing resources, time, and memory. Quantitative and qualitative assessment of the use of this approach, and in-depth analysis of the results of its application is a topic for further, more thorough research.

4. References

- [1] Elsken, T., Metzen, J. H., & Hutter, F. (2019). Neural architecture search: A survey. *The Journal of Machine Learning Research*, 20(1), 1997-2017.
- [2] Dasgupta, D. (1998). *Artificial Immune Systems and Their Applications*. Springer Berlin Heidelberg.
- [3] De Castro, Leandro & Von Zuben, Fernando. (2002). Learning and Optimization Using the Clonal Selection Principle. *Evolutionary Computation*, IEEE Transactions on. 6. 239 - 251. 10.1109/TEVC.2002.1011539.

Application of the Results of Modern Science in the Development of Information Systems for the Functioning of the Hotel Business

Maksym Holik^a, Iurii Mishcheriakov^a

^a Kharkiv National University of Radio Electronics, Nauky Ave. 14, Kharkiv, 61166, Ukraine

Abstract

Article describes the transformative impact of scientific advancements on the hotel industry, highlighting the challenges and opportunities that hotels face in harnessing these advancements to provide exceptional guest experiences, operational excellence, and environmental stewardship.

Keywords ¹

AI, Big Data, Information systems, IoT, hotel business, modern science, robotics, VR/AR

1. Introduction

In an era defined by relentless scientific discovery, the hotel industry finds itself at a transformative crossroads. The infusion of modern science into the realm of information technologies and systems has revolutionized the functioning of the hotel business. By leveraging findings and innovations from an array of scientific disciplines, hotels have achieved unprecedented levels of guest satisfaction, operational efficiency, and environmental sustainability.

2. Problem definition

The hotel industry stands at a pivotal juncture, propelled by the relentless surge of scientific discoveries that revolutionize the realm of information technologies and systems. This infusion of modern science has opened doors to unprecedented opportunities for hotels to elevate guest satisfaction, optimize operational efficiency, and embrace environmental sustainability. However, amidst these advancements, hotels face the daunting task of effectively harnessing and managing a burgeoning network of interconnected technologies, safeguarding sensitive guest data, and ensuring the seamless operation of increasingly complex systems.

In essence, the hotel industry faces a multifaceted challenge in harnessing the transformative power of modern science while navigating the complexities of data management, cybersecurity, and system integration. Addressing these challenges will be crucial for hotels to fully capitalize on scientific advancements and establish themselves as leaders in providing exceptional guest experiences, operational excellence, and environmental stewardship.

3. Ways of applying the results of science

3.1. The Internet of Things: Revolutionizing the Hotel Industry

The Internet of Things (IoT) has emerged as a linchpin in redefining the concept of smart hotels. By seamlessly interconnecting a myriad of devices and systems, IoT technology has ushered in a

Proceedings of the *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine

EMAIL: maksym.holik@nure.ua (A. 1); iurii.mishcheriakov@nure.ua (A. 2)

ORCID: 0009-0005-0317-5054 (A. 1); 0000-0002-5334-1808 (A. 2)



harmonious guest experience. Mobile applications empower guests to control room elements such as lighting and climate, while simultaneously enabling hotels to optimize resource consumption, transcending guest satisfaction into the realm of operational efficiency [1].

3.2. Big Data and Analytics: The Science of Personalized Hospitality

The utilization of Big Data and analytics has introduced a data-driven paradigm to the hotel industry. The ability to decipher guest behavior and preferences empowers hotels to personalize services, fine-tune marketing campaigns, and forecast room demand with remarkable precision [2]. The resultant enhancement of the guest experience is inextricably linked to the application of scientific insights.

3.3. Artificial Intelligence: Automating Hotel Operations for Efficiency and Guest Satisfaction

Artificial Intelligence (AI) has transcended the conventional boundaries of technology by automating a gamut of hotel operations. From expediting check-ins to optimizing room temperatures, AI systems are pivotal in enhancing operational efficiency. Chatbots, driven by AI, elevate the guest experience by addressing inquiries, and predictive AI capabilities optimize resource allocation and supply chain management.

3.4. Cybersecurity: Safeguarding Guest Data in a Digital Era

As hotels increasingly rely on digital systems, safeguarding guest data from cyber threats becomes paramount. The integration of cutting-edge cybersecurity solutions, propelled by scientific research, shields sensitive guest information and fortifies the hotel's reputation by ensuring data confidentiality and guest trust.

3.5. Environmental Sustainability: Science Driving Green Practices in Hotels

Environmental sustainability is a key focus in the hotel industry, underpinned by scientific innovation. Hotels are adopting eco-friendly technologies to reduce energy consumption, minimize waste, and lower their environmental footprint. IoT sensors optimize energy usage, while data analytics reveal opportunities for resource conservation, emphasizing the pivotal role of science in achieving sustainability [3].

3.6. Biotechnology: Enhancing Guest Health and Wellness through Scientific Innovation

The hotel industry is also embracing the potential of biotechnology to enhance guest health and wellness. Biometric sensors, embedded in wearable devices or hotel amenities, monitor guest vital signs and alert staff in case of emergencies. Biodegradable materials, derived from organic sources, replace plastic and synthetic products, reducing waste and pollution. Biotechnology also enables hotels to offer personalized wellness services, such as genetic testing, nutritional advice, and customized treatments [4].

3.7. Robotics: Transforming Hospitality with Intelligent Machines

Another scientific domain that has influenced the hotel sector is robotics. Robots, equipped with AI and IoT capabilities, perform a variety of tasks, such as delivering room service, cleaning rooms, and providing concierge services. Robots enhance guest convenience and satisfaction, while also reducing

labor costs and human errors. Moreover, robots can interact with guests in multiple languages, catering to the diverse needs of international travelers [5].

3.8. Virtual Reality and Augmented Reality: Immersive Experiences for the Modern Guest

The hotel industry is also exploring the possibilities of virtual reality (VR) and augmented reality (AR) to create immersive and engaging guest experiences. VR and AR technologies enable guests to experience different destinations, cultures, and activities, without leaving their rooms. Guests can also use VR and AR devices to access information, entertainment, and social media, enhancing their comfort and enjoyment. VR and AR technologies also provide opportunities for hotels to showcase their facilities, services, and attractions, and to create virtual tours and simulations [6].

4. Conclusions

In summation, the infusion of modern scientific knowledge into the development of information technologies and systems has indelibly impacted the hotel business. This symbiotic relationship between science and the hotel industry has ushered in an era of unparalleled guest satisfaction, operational efficiency, and environmental stewardship. As technology advances unabated, the hotel sector is poised to continue reaping the dividends of contemporary scientific discoveries.

Successfully achieves goals by highlighting the transformative impact of science on the hotel industry and providing a detailed analysis of the challenges and opportunities that hotels face in harnessing these advancements. The article serves as a valuable resource for hoteliers and hospitality professionals seeking to understand and implement scientific innovations to enhance their businesses.

5. References

- [1] Gretzel, U., Sigala, M., Xiang, Z., & Koo, C. (2015). Smart tourism: foundations and developments. *Electronic Markets*, 25(3), 179-188.
- [2] Xiang, Z., Du, Q., Ma, Y., & Fan, W. (2017). A comparative analysis of major online review platforms: implications for social media analytics in hospitality and tourism. *Tourism Management*, 58, 51-65.
- [3] Sigala, M. (2020). Tourism and COVID-19: Impacts and implications for advancing and resetting industry and research. *Journal of Business Research*, 117, 312-321.
- [4] Cohen, J., & Cohen, S. A. (2019). New directions in biotechnology and wellbeing: The case of personal genomics and leisure travel. *Journal of Travel Research*, 58(1), 79-91.
- [5] Ivanov, S., Webster, C., & Berezina, K. (2017). Adoption of robots and service automation by tourism and hospitality companies. *Revista Turismo & Desenvolvimento*, 27/28, 1501-1517.
- [6] Jung, T., tom Dieck, M. C., Lee, H., & Chung, N. (2016). Effects of virtual reality and augmented reality on visitor experiences in museum. In *Information and communication technologies in tourism 2016*, 621-635.

Development of Universal Solution for Integration with Lighting Devices

Anton Kireev^a, Artem Skriabin^a, Nataliia Serdiuk^a

^a Kharkiv National University of Radioelectronics, Nauky ave., 14, Kharkiv, 61166, Ukraine

Abstract

It is proposed to develop an innovative solution for determining the lighting color of devices depending on various factors. The main goal is to integrate artificial intelligence and provide the user with an application for selecting the lighting mode, which will allow to quickly and intuitively turn on the necessary palette in all connected lighting devices (Computer components, LED tapes, lamps, etc.) at once.

Keywords ¹

Artificial intelligence, microcontroller, optical sensors, intelligent control unit

1. Introduction

In today's world, the use of artificial intelligence (AI) is becoming a key aspect in creating smart systems, including lighting. Our project aims to develop a universal smart lighting system capable of automatically adapting to different environmental conditions and events by utilizing Bluetooth and Wi-Fi technologies present in various lighting devices.

2. Development of an innovative solution

A key feature of the system being developed will be the ability of lighting devices to change color schemes depending on the time of day, season and various events using an intelligent control unit. This unit will analyze data on time of day, season of the year and calendar holidays to dynamically adjust color scenarios. Lighting devices will change color temperatures, creating warmer hues in the evening and cooler hues in the morning hours. The system will also adapt color schemes according to the time of year, providing warmer or cooler shades depending on the current season. In addition, the lighting units will be able to automatically respond to various calendar events and holidays by changing their color palette to accentuate the festive atmosphere. Additionally, the smart control unit will take into account weather and outdoor temperature data to adapt color scenarios. In the event of cold or warm weather, the lighting devices will create the appropriate color tones, enriching the ambience of the room. Our development involves the use of a specialized color microcontroller, which will be responsible for analyzing and transmitting color data to the lighting device.

This microcontroller will be integrated into the system and will collect color information from the external environment. The color microcontroller will process data received from optical sensors or other sources (e.g. Smart Home system) capable of recording spectral characteristics of color. Using special signal processing algorithms and color models, the microcontroller analyzes the data and determines the exact color parameters or color gamut. To send color information to the lighting device, the microcontroller will use appropriate wireless communication protocols, such as Bluetooth or Wi-Fi, depending on the capabilities and settings of the particular device.

¹Proceedings of the *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine
EMAIL: anton.kiriev@nure.ua (A. 1); artem.skriabin@nure.ua (A. 2); nataliya.serdyuk@nure.ua (A. 3)
ORCID: 0000-0002-0107-4365 (A. 3)



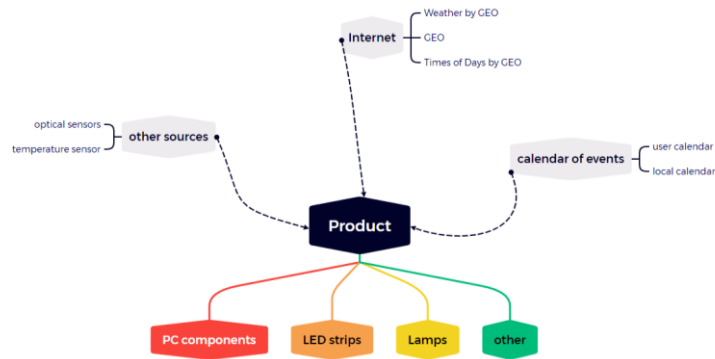


Figure 1: Schematic image of program

For example, through wireless communication, the microcontroller will send color and brightness setting commands to the lighting device. Our project develops algorithms for determining the optimal lighting color scheme, taking into account a variety of factors in the environment. To adapt to the time of day, we are developing an algorithm that automatically adjusts the color temperature according to human biological rhythms. In the evening, the algorithm will prefer warmer shades, while in the morning it will focus on cooler colors, promoting a comfortable wake-up. To adapt to the seasons, an algorithm is used that automatically changes the color palette according to seasonal changes. Summer will favor brighter colors, while winter will take into account the need for warmer shades. Integration with the user's event calendar involves developing an algorithm that can adapt the lighting to different holidays or events specified in the calendar. For example, for holidays, the system can automatically select color schemes appropriate for that event, creating a festive atmosphere. Adaptation to weather conditions will be done by using an algorithm that analyzes outdoor weather data. For example, in case of sunny weather, the system can choose brighter and lighter shades, while in cloudy weather - preference will be given to warmer and cozier colors. Each of the factors described (time of day, seasons, user's calendar of events and outdoor weather) will be presented as a separate "Lighting Mode". These modes will be available for the user to select and activate based on their preferences or current needs.

For example, the selection of the "Seasons" mode is realized by the intelligent control unit and the system will automatically switch the lighting to the mode corresponding to the current season. This means that the system will independently adapt the color palette according to the algorithms associated with the seasons with the intervention of the intelligent control unit. In this way, the intelligent control unit is able to select and switch lighting modes according to predetermined algorithms or current needs, which provides flexibility and convenience in the use of the smart lighting system.

3. Conclusion

The project will result in the development of a unique smart lighting system that not only provides functionality and comfort, but also transforms the ambience of spaces by adapting to a variety of conditions and events. This represents a significant step in the use of AI to create an adaptive and intelligent lighting system.

4. References

- [1] Philips Lighting (now "Signify"). Technology: "Philips Hue" 2012. URL: https://en.wikipedia.org/wiki/Philips_Hue
- [2] Andrew Birt, John Bosshard, Phil Bosua, Andy Gelme. Technology: "LIFX" 2012. URL: [https://en.wikipedia.org/wiki/LIFX#:~:text=LIFX%20\(pronounced%20Life%2DX\),as%20a%20smartphone%20or%20smartwatch](https://en.wikipedia.org/wiki/LIFX#:~:text=LIFX%20(pronounced%20Life%2DX),as%20a%20smartphone%20or%20smartwatch).
- [3] Tom Rodinger, Gimmy Chu, Colin West. Technology: "Nanoleaf" 2012. URL: <https://en.wikipedia.org/wiki/Nanoleaf>

Development of an intelligent visual indicator of weather conditions

Vladyslav Levchenko^a, Nikita Nechyporenko^a, Nataliia Serdiuk^a

^a *Kharkiv National University of Radioelectronics, Nauky ave., 14, Kharkiv, 61166, Ukraine*

Abstract

It is proposed to develop an innovative LED strip that automatically changes its color according to the temperature outside. The main goal is to integrate artificial intelligence and provide users with a visual indicator of weather conditions that will allow them to quickly and intuitively perceive changes in atmospheric conditions using the color palette of the LED strip.

Keywords ¹

Artificial intelligence, LED strip, microcontroller, weather indicator

1. Introduction

In 2023, artificial intelligence (AI) becomes not just a promising technology, but a relevant and necessary component for the development of various spheres of society and business, this is manifested in the optimization of business processes, where AI allows you to automate and optimize many tasks and processes in business - from analysis data and forecasting market trends for automation of production and logistics. Thanks to AI, the processing of large volumes of data has already become more accurate and efficient, which is positively reflected and helps to make informed decisions based on analysis and forecasting, which is important for business management, economy, because the execution of such tasks by a person takes a lot of time, and can cause inaccuracies in calculations. Effective marketing and advertising using the latest technology allows you to precisely direct advertising to the target audience, conduct personalized campaigns and analyze consumer behavior, reducing advertising costs, while increasing conversion and attracting customers to the product.

2. Artificial intelligence and LEDs

The OpenLicht intelligent design project, which was launched in September 2016, presents customised lighting solutions based on open source code and new materials. OpenLicht is funded by the German Federal Ministry of Education and Research to create new forms of collaboration between science, industry and manufacturers. Infineon Technologies AG is leading the project in close collaboration with Bernitz Electronics, Deggendorf Institute of Technology and Dresden University of Technology. Open Photonics projects pursue a broad spectrum of objectives. These include open innovation approaches for better utilisation of photonic components and systems as well as open source approaches for use in a wider range of applications.

¹Proceedings of the *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine
EMAIL: vladyslav.levchenko2@nure.ua (V. Levchenko); nataliya.serdyuk@nure.ua (N. Serdiuk),
nikita.nechyporenko@nure.ua (N. Nechyporenko)
ORCID: 0000-0002-0107-4365 (N. Serdiuk)



The results have been published and include a prototype for an intelligent lighting system based on artificial intelligence. It automatically adapts the lighting in a room to the user's position and activity, e.g. reading or watching TV, learns the person's preferences and can even react to new circumstances to a certain extent. The solution being developed as part of the project is based on open source approaches such as openHAB, smart home systems and machine learning libraries. The use of freely available development environments, software environments and low-cost hardware solutions enables the integration of a wide range of different sensor data and the further development of existing results by the community. Intelligent lighting control is finding its way into the smart home and will accompany us in our everyday lives with fully automated solutions. However, the solutions currently available on the market often present users with major challenges. They are often semi-automated at best, difficult to programme and operate, offer no privacy protection or consume energy unnecessarily, as the use of light is not optimally tailored to the user's needs, which in turn has a negative impact on CO2 emissions. OpenLicht offers answers to these problems. The use of artificial intelligence in the local network creates intelligent lighting solutions that are safe and sustainable and protect the privacy of users. The artificial intelligence works in a system that is shielded from the outside world and requires no connection to the internet. To achieve this, an open source gateway was developed based on the Raspberry Pi and the Trusted Platform Module (TPM) from Infineon. This means that data does not need to be sent to the cloud, but can be processed locally, ensuring the security and privacy of households. These factors are crucial for the growing popularity of smart home solutions. In addition, the automatic adaptation to the user's actions ensures that the system put into operation at a specific point in time is always ready for use.

3. Development of an innovative solution

The goal of the project is to create an innovative LED strip that automatically changes its color according to the temperature outside. The main goal is to integrate artificial intelligence and provide users with a visual indicator of weather conditions that will allow them to quickly and intuitively perceive changes in atmospheric conditions using the color palette of the LED strip. A schematic representation of the innovative LED strip is shown in Figure 1.

Sensor temperature reading is a temperature reading system that will be integrated into the LED strip, it may include the use of thermal sensors or other technologies to accurately measure the air temperature.

LED control, which will be implemented through a microcontroller, which will receive data from special software with integrated AI (intelligent unit), which will process information about the weather, and transmit it to the microcontroller of the tape, which in turn will correspond to the algorithm and ensure a smooth change colors on the LED strip.

Interaction with other devices can be a key negative of the LED strip. Interactions with other devices or systems, such as a smart home, all types of advertising - banner, interactive. This can make it possible to expand the convenience of use in everyday life, for example, when the temperature outside the window is low, the tape will create a "frosty" effect inside the house, where in fact there is a pleasant, warm and cozy atmosphere. The effectiveness of advertising such a product, which is advertised with the help of additional interactive lighting, increases many times, because it distinguishes the ad from others.

Therefore, the use of such technology can become an innovative solution. Ribbons occupy a significant place in the field of marketing, advertising, in the illumination of architectural details, bridges, building facades and other urban elements, they create impressive visual effects and change the appearance of objects, in theatrical and stage events - concerts, festivals, parties and other LED strips at entertainment events. By adding an intelligent block to the tape, you can achieve an impressive, new result, or improve an existing one.

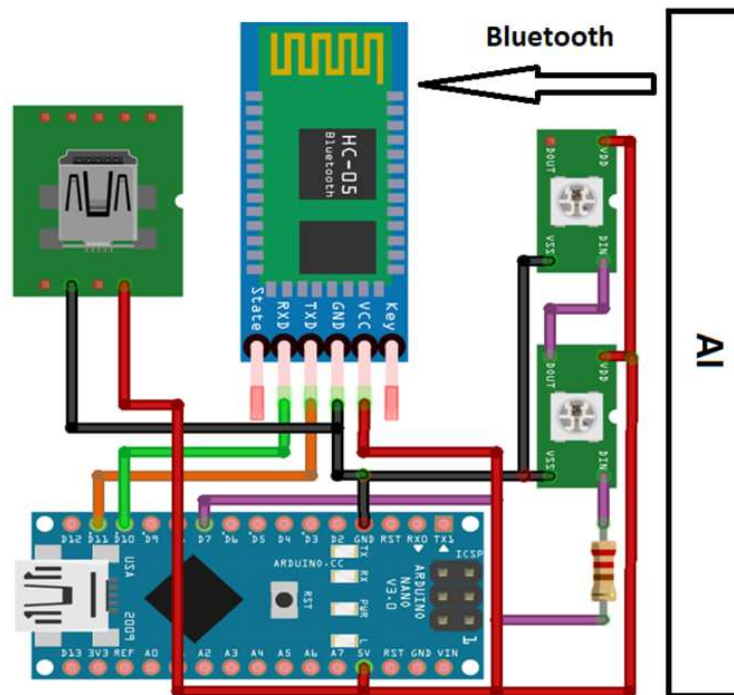


Figure 1: Schematic image of an innovative LED strip

4. Conclusion

Therefore, the use of such technology can become an innovative solution. Ribbons occupy a significant place in the field of marketing, advertising, in the illumination of architectural details, bridges, building facades and other urban elements, they create impressive visual effects and change the appearance of objects, in theatrical and stage events - concerts, festivals, parties and other LED strips at entertainment events. By adding an intelligent block to the tape, you can achieve an impressive, new result, or improve an existing one.

5. References

- [1] Bierzynski, Kay, Florian Kallerder, Pavel Lutskov, Frank Rohde, David Morales-Rodríguez, Juan Mena-Carrillo, René Schöne, and Uwe Aßmann. "OpenLicht - A Self-Learning Lighting System Based on OpenHAB," 156:43-52. SPUB 2018.
- [2] Kay, Bierzynski, Pavel Lutskov, Juan Mena-Carrillo, René Schöne, and Uwe Aßmann. "The Learning of the OpenLicht System, a Self-Learning Lighting System at the Edge of the Network," SSI 2018.
- [3] Mey, Johannes, René Schöne, Görel Hedin, Emma Söderberg, Thomas Kühn, Niklas Fors, Jesper Öqvist, and Uwe Aßmann. "Continuous Model Validation Using Reference Attribute Grammars." In Proceedings of the 11th ACM SIGPLAN International Conference on Software Language Engineering, 70-82. SLE 2018. ACM, 2018
- [4] Schöne, Rene, Johannes Mey, Boqi Ren, and Uwe Aßmann. "Bridging the Gap between Smart Home Platforms and Machine Learning Using Relational Reference Attribute Grammars." In Proceedings of the 14th International Workshop on Models@run.Time. Munich, 2019.
- [5] Juan Davila, "The Impact of Artificial Intelligence (AI) in LED Lighting" designing lighting (2021), pp. 44-45
- [6] URL: https://issuu.com/designinglighting/docs/issue_april_2021/s/12350044
- [7] Yining Chen, "Invisible Light Switch for Energy Saving Enhancement (2019)
- [8] URL:https://www.ledinside.com/press/2019/6/osram_project_invisible_light_switch_energy_saving_enhancement

Queueing Models in Cloud Computing

Olena Mikhnova^{a,1}, Danylo Obushko^{a,2}

^a Kharkiv National University of Radio Electronics, prospect Nauky, Kharkiv, 61166, Ukraine

Abstract

Recently, cloud computing has become actively deployed in IT-companies information systems due to a number of its benefits that comprise high availability, control and security. However, with this, there still remains a gap in resource distribution which can be resolved with traditional queuing. A number of such approaches are already utilized in this field, and the current research deals with analyzing the existing models, attempting at applying a combination of them to cope with customer service time and the cost of queue maintenance. Among such techniques, Mason-Wiener and Kennedy-Nishikane models have been analyzed, aiming at optimal resource distribution in a queue.

Keywords¹

Information system, Wiener model, Kennedy-Nishikane model, queueing model, cloud computing

1. Introduction

Cloud computing is an important paradigm in IT industry. It has a significant impact on modern IT companies, enabling them to reduce costs, increase flexibility and improve availability. In addition to the main characteristics, cloud computing can also be classified into the following types. Infrastructure as a Service (IaaS): provides access to virtual servers, disk storage, network resources and other IT infrastructure components. Platform as a Service (PaaS): provides access to a platform for developing, testing and deploying software. Software as a Service (SaaS): provides access to ready-to-use software products, such as email, CRM and content management systems [1].

Cloud computing is widely used in a variety of industries, including: finance, retail, healthcare, education. It is used to store financial data, process payments and provide financial services. Cloud computing is used to store customer information, manage warehouses and provide e-commerce. It is used to store medical data, diagnose diseases and provide medical services. It is used to store educational content, deliver online courses and manage educational institutions [2].

2. Queueing models in the context of cloud computing

In this section the analysis of the concept of queueing models and its role in solving the problems occurred in cloud computing are observed. Queueing models are one of the main approaches to modeling and analyzing queueing systems. They allow describing the probability that a queue will have a certain number of items at a certain point in time. For example, Mason-Wiener formula may be deployed to predict the average time to serve customers in a queue. This model assumes that customer service times are independent and distributed according to a normal distribution law.

$$W = \frac{\left(\frac{\lambda}{\mu}\right)^2}{(\mu - \lambda)} \quad (1)$$

Proceedings *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine

EMAIL: olena.mikhnova@nure.ua (A. 1); danylo.obushko@nure.ua (A. 2)

ORCID: 0000-0002-6558-8509 (A. 1); 0009-0003-3114-9959 (A. 2)



where W is an average waiting time of a request being present in a queue, μ is an average intensity of a request servicing, and λ is an average intensity of receiving the requests by the system.

Probabilistic queuing models can be used to determine the optimal queue size. This can be useful to ensure an optimal balance between the cost of queue maintenance and the loss of customer abandonment.

On the other hand, Kennedy-Nishikane formula can be used to determine the optimal queue size. This model assumes that the arrival of customers is independent and distributed according to the Poisson law, and the time of customer service is independent and distributed according to the exponential distribution law.

$$W = \frac{L}{\mu} \quad (2)$$

where W is an average waiting time of a request being present in a queue, μ is an average intensity of a request servicing, and L is an average number of requests in the system.

The parametric approximation model provided further can be implemented for IS functioning description, which is a nonlinear system, in order to formalize the process of queue formation in cloud computing. Let an input sequence of a system be presented as $u(t)$, an output sequence as $y(t)$, an intermediary state as $x(t)$, and $w(t)$ as outliers present in intermediary states. The input sequence may be introduced as $u^N = \{u(1), u(2), \dots, u(N)\}$, and the probability function of the output as $y^N = \{y(1), y(2), \dots, y(N)\}$. Wiener model for such a nonlinear system can be represented by the following system of equations:

$$\begin{cases} x_0(t) = G(q, \theta) u(t) \\ x(t) = x_0(t) + w(t) y(t) \end{cases} \quad (3)$$

where q and θ are the estimated parameters that match initial data sequence. As referred to [3, 4], a model class for G may be introduced in a form of FIR filters, rational transfer functions (OE models), state space models (which are the matter in our case), Laguerre filters.

3. Conclusion

In addition to the tasks discussed, queueing models can also be used for queue maintenance cost reduction. To achieve this goal, we have analyzed existing queueing models for cloud computing. This task is necessary in order to identify the strengths and weaknesses of existing models, as well as to identify opportunities for their usage and improvement. The analysis allowed us to combine the most efficient models to our particular case. A combined queueing model has been tested for compliance with the requirements based on one of the market leading IT companies located in Kharkiv.

As a result of the work, an analysis of existing research in the field of cloud computing and queueing models was carried out. Also the available statistical data on the load on the IS of an IT company has been analyzed, the requirements for the quality of customer service have been considered. Thus, the performed study aims at developing and optimization of a queueing model to improve the operation of IT companies ISs.

4. References

- [1] S. Orban, Ahead in the Cloud: Best Practices for Navigating the Future of Enterprise IT (1-st ed.), Create Space Independent Publishing Platform, 2018.
- [2] A. Bahga, V. Madiseti, Cloud Computing: A Hands-On Approach, CreateSpace Independent Publishing Platform, 2013, 454 p.
- [3] A. Hagenblad, L. Ljung, A. Wills, Maximum likelihood identification of Wiener models, Automatica 44 (2008) 2697-2705. doi:10.1016/j.automatica.2008.02.016.
- [4] L. Sersour, T. Djamah, M. Bettayeb, Nonlinear system identification of fractional Wiener models, Nonlinear Dynamics 92 (2018) 1493–1505.

Investigation of the Efficiency Dependence of Relational and Graph Databases on Data

Ihor Nedaiev^a, Oksana Mazurova^a and Marija Shirokopetleva^a

^a Kharkiv National University of Radio Electronics, 14 Nauky Avenue, Kharkiv, 61166, Ukraine

Abstract

The purpose of the study is to analyze the methods of working with relational and graph databases, to develop recommendations for their use depending on the input data.

Keywords ¹

Data, dependence, efficiency, graph database, metric, relational database

1. Introduction

Nowadays, databases are an integral part of most application software systems under development. During development, different types of databases are used, where the logic of each of them has its own purpose, strengths and weaknesses. Accordingly, each type of DBMS (database management system) has its own specialized field of use and was created to solve its own range of tasks. In order to model complex, loosely structured subject areas with different types of data, relationships, etc., developers face an important task of determining which type of DBMS is best suited in this context.

2. Problem statement

The purpose of this paper is to formulate more effective recommendations for choosing a data model and a suitable DBMS taking into account the context of the data that will form the basis of the respective databases based on the analysis of the results of their experimental studies.

To achieve this goal, the following tasks need to be solved:

- to select the best database systems for further research based on the analysis of existing DBMSs and multi-criteria decision-making;
- plan an experimental study (select a subject area and design a database, develop a query system for the study and select metrics for the study);
- develop software solutions for each of the studied DBMSs;
- conduct experiments and formulate recommendations for choosing a DBMS in accordance with the context of the subject area data.

3. Description of the study

It was decided to focus on the study of the efficiency of relational and graph [1,2] databases and their dependence on the data context. The object of the study is the process of working with graph and relational databases, and the subject of the study is the methods of storing, processing and accessing data in relational and graph database management systems, in particular, lazy loading using the Entity Framework.

There are some general recommendations in the literature on choosing a database, but there is no clear division - the choice is up to the developers. Many studies show that graph databases are much

Proceedings of the *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine

EMAIL: ihor.nedaiev@nure.ua (A. 1); oksana.mazurova@nure.ua (A. 2); marija.shirokopetleva@nure.ua (A. 3)

ORCID: 0009-0000-6194-8143 (A. 1); 0000-0003-3715-3476 (A. 2); 0000-0002-7472-6045 (A. 3)



more efficient than relational databases if the database contains many relationships between tables [2], while with a large amount of data and a small number of relationships, their performance is much lower [1, 3].

We analyzed such popular graph DBMSs as: Neo4j, ArangoDB, OrientDB-X, Amazon Neptune, and Azure Cosmos DB. In order to select a DBMS for further research, the task of multicriteria selection was solved, taking into account such criteria as response time, scalability, data consistency and support for ACID properties during transaction processing, support for a graphical data model, ease of use and development.

The qualitative assessments obtained during the analysis were converted to quantitative scales, normalized, and analyzed using the Pareto principle. A convolutional model based on linear additive convolution with weighting coefficients was chosen for the multicriteria selection:

Based on the results of solving the multicriteria decision-making problem, the Neo4j graph DBMS was chosen for further research, since Amazon Neptune is focused on larger projects and has limited capabilities when used for free.

As a subject area for modeling and studying the dependence on the data context, the field of football statistics was chosen, which is described by data of various types, including media data, and requires modeling various types of relationships, for example, the history of encounters between two players, etc.

Researching the process of working with graph and relational databases involves studying the ways of storing, processing, and accessing data in these systems, as well as their differences in terms of efficiency.

As methods of storing, processing and accessing data in relational and graph database management systems, it was decided to consider the option of lazy loading using the Entity Framework as well. Lazy loading is a technique used to postpone the loading of related data until the moment when it is really needed.

4. Conclusion

At this stage, the most suitable database systems were selected for further research and an experimental study was planned. In the future, the obtained results will be used to write a master's thesis, in particular, to develop a software solution and develop recommendations for the choice of a DBMS depending on the subject area based on the experiments.

5. References

- [1] Kotiranta P., Junkkari M., Nummenmaa J. Performance of graph and relational databases in complex queries. *Applied sciences*. 2022. Т. 12, № 13. С. 6490. URL: <https://doi.org/10.3390/app12136490> (дата звернення: 03.04.2023).
- [2] Lazarska M., Siedlecka-Lamch O. Comparative study of relational and graph databases. 2019 IEEE 15th international scientific conference on informatics, м. Poprad, Slovakia, 20–22 листоп. 2019 р. 2019. URL: <https://doi.org/10.1109/informatics47936.2019.9119303> (дата звернення: 03.04.2023).
- [3] Mazurova, O. NOSQL database logic design methods for MONGODB and NEO4J / Mazurova, O., Syvolovskyi, I., Syvolovska, O. *Innovative technologies and scientific solutions for industries*, № 2 (20), pp. 52-63.

Simulation of the page replacement process

Hanna Novakova^a, Nataliia Khatsko^{a,b} and Kyrylo Khatsko^a

^a National Technical University «Kharkiv Polytechnic Institute», Kharkiv, 61002, Ukraine

^b National University of Radio Electronics, Kharkiv, 61166, Ukraine

Abstract

An efficient page replacement algorithm is required by the operating system for the efficient execution of applications. FIFO, Second Chance, etc. are used to select the page to be unloaded in case of a page fault. Many algorithms have been developed over the years. Each of them aims to reduce the number of page faults. This article presents the results of experiments of the replacement process with the FIFO algorithm. The occurrence of Belady's anomaly is demonstrated. A proposal for improving modern algorithms is given – the introduction of a two-bit variable to determine page activity.

Keywords ¹

Page replacement, Belady's anomaly, FIFO, Second Chance, physical memory.

1. Introduction

Let us consider the process of page swapping in main memory, a problem that has been studied for a long time [1, 2, 3, 6, 7]. It is known that RAM is divided into page frames that have the same constant size. When preparing for execution, the process is placed in virtual memory, which is also divided into pages of the same size as the frames. Before executing the process, the virtual pages of the process must be loaded into free RAM frames. If there is no free space, the operating system will free up the necessary space to place the page.

According to the demand paging, the execution of the process starts with empty physical memory, the page needs to be loaded into main memory only when it is requested. If the requested page is not present in physical memory, a page fault is said to occur. Based on the scheme described above, the number of loads affects the performance of the computer system, so it is important to reduce the number of loads. There are many algorithms for replacing pages in RAM (FIFO, LRU, Second Chance). For the study, we used FIFO, the simplest of the algorithms.

2. Results of the experiment

FIFO is simple to understand and implement a page replacement algorithm. It removes the page that was loaded into physical memory first. A linked list is usually used for this page replacement. If a page fault occurs, it deletes one at the head and attaches a new page to the tail of the list. The performance of this algorithm is rather bad because it suffers from Belady's anomaly [3].

Belady's anomaly is unexpected behavior in FIFO. In some of the reference strings, increasing the number of page frames increases the number of page faults [4].

For a particular application, a sequence of requests for the pages of process that are defined in virtual memory has been determined. For the experiment, a string containing this sequence was created as an input.

Proceedings of the *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine
EMAIL: hanna.novakova@cs.khpi.edu.ua (H. Novakova); nataliia.khatsko@khpi.edu.ua (N. Khatsko); kyrylo.khatsko@khpi.edu.ua (K. Khatsko)

ORCID: 0009-0006-8115-431X (H. Novakova); 0000-0002-2543-0280 (N. Khatsko); XXXX-XXXX-XXXX-XXXX (K. Khatsko)



For example, consider the reference string that has numbers from 1 to 6. The sequence is defined in the first row of each table from Tables 1-4. The first column lists the page frames in RAM, and their number corresponds to the conditions of the experiment.

The rows correlate with the page frames containing the numbers of virtual pages loaded into RAM. In the line Faults: 1 – there is no page in the RAM and the process needs to load one; 0 – the page is already loaded. The last line shows the resulting number of required replacements.

In Tables 1-4, you can see the process of replacing pages with the FIFO algorithm.

Table 1

The process of replacing pages with a FIFO algorithm with three frames for distribution

Pages	1	2	1	3	4	1	6	2	4	3	4	2
Frame 1	1	1	1	1	4	4	4	2	2	2	2	2
Frame 2	-	2	2	2	2	1	1	1	4	4	4	4
Frame 3	-	-	-	3	3	3	6	6	6	3	3	3
Faults	1	1	0	1	1	1	1	1	1	1	0	0

Sum of faults – 9

Table 2

The process of replacing pages with a FIFO algorithm with four frames for distribution

Pages	1	2	1	3	4	1	6	2	4	3	4	2
Frame 1	1	1	1	1	1	1	6	6	6	6	6	6
Frame 2	-	2	2	2	2	2	2	2	2	2	2	2
Frame 3	-	-	-	3	3	3	3	3	3	3	3	3
Frame 4	-	-	-	-	4	4	4	4	4	4	4	4
Faults	1	1	0	1	1	0	1	0	0	0	0	0

Sum of faults – 5

Table 3

The process of replacing pages with a FIFO algorithm with three frames for distribution

Pages	1	2	3	4	1	2	5	1	2	3	4	5
Frame 1	1	1	1	2	3	4	1	1	1	2	5	5
Frame 2	-	2	2	3	4	1	2	2	2	5	3	3
Frame 3	-	-	3	4	1	2	5	5	5	3	4	4
Faults	1	1	1	1	1	1	1	0	0	1	1	0

Sum of faults – 9

Table 4

The process of replacing pages with a FIFO algorithm with three frames for distribution

Pages	1	2	3	4	1	2	5	1	2	3	4	5
Frame 1	1	1	1	1	1	1	2	3	4	5	1	2
Frame 2	-	2	2	2	2	2	3	4	5	1	2	3
Frame 3	-	-	3	3	3	3	4	5	1	2	3	4
Frame 4	-	-	-	4	4	4	5	1	2	3	4	5
Faults	1	1	1	1	0	0	1	1	1	1	1	1

Sum of faults – 10

The experiment revealed an increase in the number of page faults with an increase in the number of RAM frames allocated to the process, which is contrary to expectations – with an increase in the number of frames, the number of loads should have decreased. Thus, we observe the well-known Belady's

anomaly. In modern operating systems, this anomaly is overcome by a number of algorithms. But the problem still exists, and everyone can observe it at some points in the computer's operation.

3. Conclusions

This paper presents the results of a computer simulation to reproduce Belady's anomaly. It is determined that it occurs when the page needs to be loaded after a short period of time.

For the FIFO algorithm, it was found that the phenomenon may or may not occur during an operation. It makes sense to characterize each process page by the degree of its importance for the process execution. To determine whether the process has accessed a particular page, a reference bit was created with which this idea can be implemented. The well-known Second Chance algorithm uses this bit, thereby allowing active pages to remain in RAM for some time. According to the results of experiments, it turned out that this implementation is not the best. We propose increasing the length of the reference variable to two bits, which can improve the characteristics of the activity of each page (the number of accesses). This way, we can distinguish between the degree of page activity and determine whether to leave the page in RAM.

In future research, we are going to modify the Second Chance algorithm and implement a two-bit characteristic of page access activity with a software.

4. References

- [1] H. Paulson, R. Ramachandran, Page Replacement Algorithms – Challenges and Trends. *International Journal of Computer & Mathematical Sciences* 6(9) (2018) URL: <https://ssrn.com/abstract=3128697>
- [2] P. M. Chawan, A Comparison of Page Replacement Algorithms. *International Journal of Engineering and Technology* 3 (2011) doi:10.7763/IJET.2011.V3.218.
- [3] G. Rexha, E. Elmazi, I. Tafa, A Comparison of Three Page Replacement Algorithms: FIFO, LRU and Optimal. *Academic Journal of Interdisciplinary Studies* 4 (2015) doi:10.5901/ajis.2015.v4n2s2p56.
- [4] K. McMaster, S. Sambasivam, N. Anderson, How Anomalous Is Belady's Anomaly? *Issues in Informing Science and Information Technology* 9 (2009) doi:10.28945/3390.
- [5] K. McMaster, S. Sambasivam, N. Anderson, A Probability Model for Belady's Anomaly. *Information Systems Education Journal* 8(43) (2010) URL: <http://isedj.org/8/43/>
- [6] T. D. Putra, R. Purnomo, P. Wowon, Analysis of Random Page Replacement Algorithm in Operating System. *IJARCCCE*. 9 (2020) 52-55 doi:10.17148/IJARCCCE.2020.91009.
- [7] S. Shastri, A. Sharma, V. Mansotra, Study of Page Replacement Algorithms and their analysis with C#. *The International Journal Of Engineering And Science (IJES)* 5(1) (2018) 53-57. URL: https://www.researchgate.net/publication/326331469_Study_of_Page_Replacement_Algorithms_and_their_analysis_with_C
- [8] P. Fornai, A. Iványi, FIFO anomaly is unbounded (2010). URL: <https://arxiv.org/abs/1003.1336>
- [9] M. A. Bender et al. APPENDIX : Cache-Adaptive Analysis (Full Version) (2017). URL: <https://research.vmware.com/files/attachments/0/0/0/0/3/9/ca-full.pdf>
- [10] A. Lincoln, Q. Liu, J. Lynch, H. Xu, Cache-Adaptive Exploration: Experimental Results and Scan-Hiding for Adaptivity. In: *Proceedings of the 30th ACM Symposium on Parallelism in Algorithms and Architectures* (2018) pp. 213-222 doi: 10.1145/3210377.3210382.

Cloud services: Challenges and Threats

Mariia Tarasenko^a, Nataliia Serdiuk^a

^a Kharkiv National University of Radioelectronics, Nauky ave., 14, Kharkiv, 61166, Ukraine

Abstract

This paper examines the common and specific challenges and threats that arise when using cloud service models. The main objective is to find solutions to these problems that can help optimise and operate the systems efficiently.

Keywords ¹

Cloud services, IaaS, PaaS, SaaS, SLA, XaaS

1. Introduction

Cloud services play a crucial role in the contemporary information and technology landscape, providing a diverse range of offerings like Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS).

The concept of Everything as a Service (XaaS) further amplifies organizational capabilities, granting unmatched flexibility and efficiency in the utilization of computing resources and software solutions. This, in turn, enhances operational capabilities and technological agility for businesses[1].

Nevertheless, as organizations progressively embrace cloud technologies, they are confronted with a series of challenges. Addressing these challenges necessitates strategic solutions to ensure the seamless integration and secure utilization of cloud services.

2. Common challenges of using cloud services

While cloud services have achieved significant success, they come with inherent challenges. These challenges encompass areas such as data security, confidentiality, network reliability, and the absence of standardized practices and interoperability among diverse cloud platforms[2]. These factors can pose difficulties in tasks like migrating, integrating, and sharing data across various cloud services.

3. Specific cloud services – AIaaS

In addition to fundamental cloud services, specialized offerings are emerging, such as Artificial Intelligence as a Service (AIaaS). AIaaS presents opportunities to leverage intelligent algorithms, machine learning, and data analysis through cloud computing[3].

AIaaS simplifies the development, training, and deployment of artificial intelligence models by providing a flexible cloud infrastructure. This enables users to access potent machine learning tools without the need for substantial investments in dedicated computing infrastructure.

Given the myriad use cases for AI, businesses often encounter challenges in constructing and maintaining a bespoke AI tool for each scenario[4]. The adaptability of AIaaS is particularly valuable as organizations can swiftly deploy AI services and tailor them to align with their specific needs.

¹Proceedings of the *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine

EMAIL: mariia.tarasenko@nure.ua (A. 1); nataliya.serdyuk@nure.ua (A. 2)

ORCID: 0000-0002-0107-4365 (A. 2)



4. Specific AIaaS problems

As the popularity of AI as a Service (AIaaS) increases, new challenges emerge that have the potential to impact the security and effectiveness of the service. These challenges may encompass issues common to cloud services in general, such as concerns about data security and vulnerabilities to unauthorized access.

Additionally, there are risks specific to artificial intelligence, like blended attacks targeting machine learning models. The integration of AIaaS into existing systems can prove intricate, particularly when confronted with diverse technology stacks and varied data formats[5].

The resource-intensive nature of AI model training and deployment can result in scalability challenges and elevated operational costs. Moreover, there exists a shortage of skilled professionals proficient in both AI and the specific domain of application, further complicating the effective implementation and management of AIaaS

5. Solving the problems of cloud services

One potential strategy to address these challenges includes strengthening encryption and data protection mechanisms, implementing advanced cybersecurity measures, refining authentication, and authorization systems, and advancing technologies related to vulnerability detection systems. Through these efforts, organizations can reduce the risks associated with the use of intelligent algorithms, enhancing the resilience of models against unauthorized interference and overall cybersecurity[2].

To ensure operational continuity and minimize downtime, organizations may opt for a multi-cloud or hybrid cloud strategy, providing redundancy. Additionally, selecting cloud providers with robust Service Level Agreements (SLAs) that guarantee high availability contributes to a more reliable infrastructure.

6. Conclusion

Solving the challenges associated with cloud services requires a multifaceted approach that encompasses technological advancements, strategic planning, and ongoing commitment to security and compliance. As organizations continue to harness the power of the cloud, adopting these solutions will pave the way for a more secure, efficient, and resilient cloud computing landscape.

AIaaS opens new opportunities for developers and researchers, but its use comes with specific challenges and threats. Its shortcomings indicate that there is still room for improvement, but despite the potential obstacles to its development, understanding these issues and improving security measures will allow AIaaS to maximise the benefits of AIaaS in various fields and this cloud service model will be as effective and important as its predecessors.

7. References

- [1] International Journal of Computer Science & Information Technology (IJCSIT) Vol 5, No 3, 2013.
- [2] "Cloud Security Risks and Solutions," White Paper, BalaBit IT Security, 2010.
- [3] L. Sharma and P. K. Garg, 'Artificial intelligence: Technologies, applications, and challenges,' 2021.
- [4] S. Lins, K. D. Pandl, H. Teigeler, S. Thiebes, C. Bayer and A. Sunyaev, 'Artificial intelligence as a service: Classification and research directions,' Business & Information Systems Engineering, vol. 63, 2021.
- [5] S. Truex, L. Liu, M. E. Gursoy, L. Yu and W. Wei, 'Demystifying membership inference attacks in machine learning as a service,' IEEE Transactions on Services Computing, vol. 14, no. 6, 2019.

Model of Threat Analysis and Detection System in Computer Networks Based on the Danger Theory of Artificial Immune Systems

Oleg Torubara^a, Nataliia Serdiuk^a

^a Kharkiv National University of Radioelectronics, Nauky ave., 14, Kharkiv, 61166, Ukraine

Abstract

The aim is to investigate a model of a system suitable for analyzing network traffic that would meet the requirements of modern cybersecurity. Due to the growing popularity and accessibility of artificial intelligence, its models have begun to be used to crack cybersecurity systems.

Keywords ¹

Artificial immune system, traffic, networks, cybersecurity

1. Introduction

Currently, the most common software for preventing network intrusions is a manually configured firewall. On the one hand, this is a classic approach, well-known and relatively reliable. Firewalls are mostly configured manually, and they can simply block or allow certain traffic according to the configuration. On the other hand, we have a growing demand for next-generation firewall (NGFW). These not only encapsulate the functionality provided by a classic firewall, but also provide new features with built-in intrusion prevention/detection systems (IPS/IDS). These solutions follow the classical approach to information systems, i.e., they follow rigidly prescribed algorithms, are not capable of learning and adaptation.

2. Threats detection methods

This paper discusses the topic of an alternative approach to analyzing network traffic using a system based on the danger theory[2] of artificial immune systems(AIS)[1]. AISs are based on the principles of the immune system of biological organisms and can be used to detect anomalies and other unusual changes in network traffic, which may be a sign of potential threats to network security. Using this approach, it is possible not only to provide a wide set of rules that determine a particular system behavior. But it is also possible to extend the existing functionality by adding an advanced layer of analysis software which will be based on artificial immune networks to perform analysis of traffic, host behavior in the network, availability of responses to requests, etc. **Figure 1.**

The idea is to perform traffic/network activity analysis in response to anomalous behavior, detecting incidents that may indicate attacks or other dangers. This includes analyzing traffic, identifying a large number of unusual requests, excessive attempts to access the system, and other anomalous patterns. The algorithms used in AIS[4] can detect these anomalies and notify network administrators or trigger automatic measures to prevent potential threats. Of course, the effectiveness of the system depends on the proper training and configuration of the AIS, as well as on the ability to detect new threats. This technology can be useful for improving the security of computer networks and

¹Proceedings of the *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine
EMAIL: oleg.torubara@nure.ua (O. Torubara); nataliya.serdyuk@nure.ua (N. Serdiuk)
ORCID: 0000-0002-0107-4365 (N. Serdiuk)



detecting potential attacks. The diagram shows the traffic path through the system divided into three modules. The first stage of the system is traffic capture at the interface, followed by primary data processing and verification for compliance with security rules configured in the system. After that, the data can be transferred to the traffic analysis module based on an AIS for more detailed traffic analysis. In this diagram, the solid lines indicate the path that the traffic passes through as a matter of course, and the dashed lines indicate the paths that are optional, which means that the data flow may be before this stage. This approach does not exclude the classic way of filtering network traffic, defining rules and analyzing traffic, but aims to add an additional level of analysis.

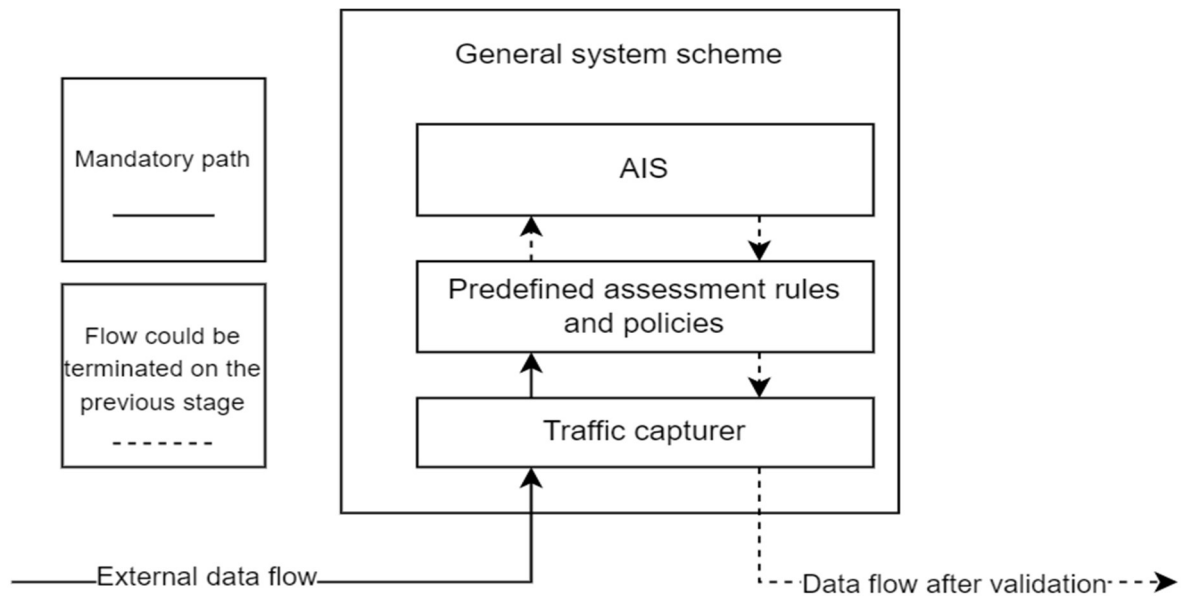


Figure 1: General scheme of traffic processing

The first approach to build such system is aimed at detecting anomalies, this approach is based on the use of a positive selection algorithm[2], whereby the system adapts to legitimate patterns of behavior in the network and detects anomalies/abnormalities. The opposite[3] approach according to which the system should detect attack signatures is based on the use of a negative selection algorithm[2]. The system adapts to identify known attack signatures and compares them with network traffic. If a match is found with known attack signatures, it may indicate an attack. An alternative is a hybrid approach, which is the use of a combination of different methods to increase the effectiveness of threat detection and make the system more flexible.

3. Conclusion

Building and configuring such a system can be a complex task and requires in-depth knowledge of network security and machine learning. Another important factor is the need to regularly update the system, as new threats and attacks are constantly changing.

4. References

- [1] L.N. de Castro and J. Timmis Artificial Immune Systems: A New Computational Intelligence Approach
- [2] U. Aickelin, S. Cayzer The Danger Theory and Its Application to Artificial Immune Systems
- [3] D.Dasgupta A comparison of negative and positive selection algorithms in novel pattern detection.
- [4] Dasgupta D., Yu S., Nino F. Recent Advanced in Artifical Immune Systems: Models and Applications

Аналіз Методів Виявлення Аномальних Значень у Наборах Експериментальних Даних

Катерина Горішня^а, Володимир Кобзев^а

^а Харківський національний університет радіоелектроніки, пр. Науки, 14, м. Харків, 61166, Україна

Анотація

Розглянуті особливості існуючих методів виокремлення аномальних значень, які застосовуються для виявлення прихованих закономірностей у наборах даних.

Ключові слова ¹

Аналіз даних, аномальні значення, методи виявлення

1. Вступ

Статистичний аналіз даних експериментальних досліджень передбачає обробку величезного обсягу інформації для виявлення існуючих в них закономірностей. Існує також проблема виявлення аномальних, нетипових даних, які можуть спотворити результати аналізу та прогнозування. В будь-якому разі, необхідно вміти виокремлювати з наявних наборів експериментальних даних такі нетипові значення, як помилкові або як можливі прояви рідких явищ, не врахованих обраною моделлю даних. В даній роботі наведені результати узагальненого аналізу найпопулярніших методів виявлення аномальних значень.

Зазвичай аномальним значенням (викидом) прийнято вважати елемент сукупності певного об'єму, що значно відрізняється набором значень своїх характеристик від інших її елементів [1]. Однак термін «значна відмінність» не завжди має точного визначення, що робить бажаним наявність кількісної міри відповідності між елементами аналізованого набору даних та характеристиками їхнього статистичного розподілу, передбачуваними обраною моделлю.

2. Методи виявлення аномальних значень

Найпопулярніші методи виявлення аномальних значень (аномалій) необхідно розглядати у відповідності до окремих задач виявлення прихованих закономірностей (задач Data Mining) [2].

1. Задача класифікації.

Припускається, що нормальна поведінка деякої системи визначатися одним чи кількома класами її характеристик. Екземпляр, який не належить до жодного з класів, є аномалією. Пошук аномалій має два етапи: навчання та розпізнавання. Класифікатор навчається на масиві маркованих даних. Визначається приналежність нових елементів до одного з відомих класів. Екземпляри, що не віднесені до жодного класу, позначаються як аномалії.

Представниками цієї групи методів розпізнавання аномалій є: нейронні мережі, байсові мережі, метод опорних векторів і метод на основі правил, які відповідають нормальній поведінці системи.

2. Задача кластеризації.

Виявлення аномалій будується на припущенні, що нормальні екземпляри утворюють великі щільні кластери, а аномальні – маленькі та розрізнені. Одним з найпростіших реалізацій підходу з урахуванням кластеризації є алгоритм k-means.

Proceedings of the *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine

EMAIL: kateryna.horishnia@nure.ua (A. 1); volodymyr.kobziev@nure.ua (A. 2)

ORCID: 0000-0002-9955-4710 (A. 1); 0000-0002-8303-1595 (A. 2)



3. Розширений статистичний аналіз.

Досліджується певна система, будується її профіль (модель), який потім порівнюється з реальними її проявами. Якщо різниця у реальній і модельній поведінці системи вище за встановлений поріг, робиться висновок про наявність аномалій. Найчастіше припускають, що нормальна поведінка системи перебуватиме в зоні високої ймовірності, тоді як аномалії у зоні низької. Цей клас методів зручний, він не вимагає попередньо певних знань про вид аномалії. Складнощі пов'язані з визначенням точного статистичного розподілу ознак та порогу.

Методи статистичного аналізу поділяються на дві основні групи:

- параметричні методи припускають, що нормальні дані мають розподіл з кінцевою множиною параметрів $P(x, \theta_i)$, де $i = 1, 2, \dots, n$. Ці методи часто ґрунтуються на Гаусовій чи регресійній моделі, а також їх комбінації;
- непараметричні методи передбачають, що структура моделі не визначена апіорно кінцевою множиною параметрів, натомість вона певним чином визначається з наявних даних. Серед найбільш відомих методів цієї групи методи з урахуванням гістограм чи функцій ядра.

Наприклад, групу найменших значень (див. рис. 1) можна віднести до аномальних.

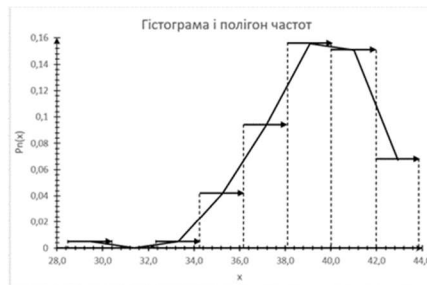


Рисунок 1: Приклад гістограми з аномаліями

Розпізнавання аномалій на основі функції ядра відбувається аналогічно до параметричних методів за винятком способу оцінки щільності ймовірності.

4. Алгоритм найближчого сусіда використовує поняття відстані між об'єктами (наприклад, Евклідова відстань). Два основних підходи базуються на таких припущеннях:

- відстань до найближчого сусіда визначається для кожного екземпляра класу, що тестується. Екземпляр, найбільш віддалений від найближчого сусіда, вважається аномалією;
- використовується відносна густина у околиці кожного екземпляра даних. Примірник, що знаходиться в проміжку з низькою щільністю, рахується аномальним. Його відстань до k-го найближчого сусіда еквівалентна радіусу гіперсфери з центром в даному екземплярі.

5. Гібридні методики розпізнавання аномалій дозволяють поєднувати переваги різних вище зазначених підходів.

3. Висновки

Таким чином, методи виявлення аномальних значень вирізняються своїми підходами та перевагами. Комплексне використання різних методів може забезпечити більш ефективне виявлення та обробку аномалій в обсягах великих даних. Зокрема, важливо враховувати особливості конкретного завдання та особливості даних для вибору оптимального методу виявлення аномальних значень.

4. Література

- [1] L Kirichenko, V Kobziev, Y Fedorenko. Data Mining methods for detection of collective anomalies in time series, In: Intern. sc. and pract. conf. "Application of information technologies in the preparation and operation of law enforcement forces", Kharkiv, 2021, p. 106.
- [2] Han, Jiawei. Data mining: concepts and techniques / Jiawei Han, Micheline Kamber, Jian Pei. – 3rd ed., Morgan Kaufmann Publishers is an imprint of Elsevier, 2012, 740p.

Дослідження Застосування Методів Аналізу Даних у Системах Автоматизації Роботи Кінотеатрів

Кирило Грач^а та Сергій Тітов^а

^а Харківський Національний Університет Радіоелектроніки, 14, просп. Науки., Харків, 61166, Україна

Анотація

У роботі розглянуто і досліджено методи аналізу даних та їх використання в розважальній індустрії, а саме у кінотеатрах. Для дослідження було створено прототип інформаційної системи бронювання квитків кінотеатру. Застосовуючи розроблений прототип системи бронювання квитків кінотеатру, виявлено вхідні та вихідні дані. На основі цих даних було досліджено три методи аналізу даних. Після дослідження результатів це дозволило підвищити ефективність управління кінотеатром.

Ключові слова ¹

Аналіз споживчої поведінки, електронна комерція, кінотеатр, метод ABC-аналізу, методи аналізу даних, часові ряди.

1. Вступ

Кінематографія - це не лише мистецтво, але й велика частина сучасної розважальної індустрії. З кожним роком кінотеатри пропонують ширший вибір фільмів та першокласні враження від перегляду. Однак цей швидкий розвиток супроводжується зростанням популярності, і, як наслідок, збільшенням кількості клієнтів і складністю керування квитками. Відповідно, інформаційні системи для бронювання квитків стають невід'ємною частиною цієї галузі, що дозволяє як клієнтам, так і кінотеатрам зручно та ефективно бронювати та керувати квитками на фільми.

В сучасному інформаційному суспільстві, де обсяги даних ростуть експоненційно, а отримання інформації стає легкодоступним завдяки широкому розповсюдженню цифрових технологій, аналіз даних стає ключовою складовою успіху в багатьох сферах життя [1]. Від бізнесу і науки до медицини і громадської політики, аналіз даних дозволяє отримувати глибокі інсайти, приймати обґрунтовані рішення та оптимізувати процеси.

2. Хід роботи

Існують різноманітні методи аналізу даних, які можна використовувати для кінотеатрів, а саме: аналіз часових рядів [2]; кореляційний аналіз [3]; аналіз впливу та експерименти [4]; геоінформаційний аналіз [5] та інші. Щоб провести дослідження методів аналізу даних потрібно знати, які вхідні та вихідні дані використовуються у системі кінотеатру. Для цього треба створити прототип інформаційної системи бронювання квитків кінотеатру. Створення прототипу складається з двох етапів: формування вимог та моделювання програмного забезпечення.

Формування вимог включає визначення ролі користувачів і їх можливості у системі, сценарії поведінки та схеми навігації. По-перше, для розробки було обрано написання користувацького сценарію у вигляді розповіді користувача про послідовність своїх дій для

Proceedings of the *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine

EMAIL: kyrylo.hrach@nure.ua (A. 1); serhii.titov@nure.ua (A. 2)

ORCID: 0009-0009-2824-8677 (A. 1); 0000-0003-0910-4415 (A. 2)



досягнення результату. На другому етапі обрано User Story як спосіб формалізованого опису дій користувача. На третьому етапі сценарії поведінки представляються у вигляді діаграми Use Case, яка описує всі взаємодії системи з користувачем. На четвертому етапі збираються разом усі сценарії, User Story і Use Case. Після чого створюється User Flow, тобто очікуваний алгоритм дій користувача, і отримується повний опис кожної взаємодії між користувачем і продуктом, а також чітке бачення основних принципів роботи продукту [6].

Моделювання програмного забезпечення складається з двох частин: клієнтської і серверної. Клієнтська частина потрібна для того, щоб точніше зрозуміти як повинна працювати система. Для цього використовується опис функціональності системи завдяки діаграмі послідовності, діаграмі діяльності та діаграмі потоків даних. Для розробки було використано StarUML, BPwin і Erwin. Для серверної частини були використані інструменти для моделювання баз даних та серверної логіки, такі як MySQL Workbench [7].

Використовуючи створений прототип було досліджено три метода аналізу даних, а також розглянуто як вони впливають на предметну область, а саме:

1. часові рядки (ARIMA). Реалізація прогнозування продажів квитків за допомогою методу ARIMA підкреслила його корисність для планування афіші та каталогу сеансів;
2. метод ABC для інвентаризації. Класифікація товарів на категорії за методом ABC дозволяє оптимізувати управління запасами, приділяючи увагу найбільш важливим та стратегічним товарам;
3. аналіз споживчої поведінки на акції. Розроблений метод аналізу споживчої поведінки при акціях дозволяє ефективно адаптувати маркетингові стратегії та оптимізувати акції для збільшення попиту та прибутку.

3. Висновки

У результаті дослідження застосування методів аналізу даних у системах автоматизації роботи кінотеатру виявилось, що використання інтегрованих підходів може значно поліпшити ефективність та стратегії управління. Застосування методів моделювання (IDEF0, DFD, BPwin, ERwin) дозволило створити детальний прототип системи кінотеатру, що дало можливість вибрати будь-який метод аналізу даних та дослідити його в предметній області.

Застосування отриманих результатів у реальному бізнес-середовищі може призвести до оптимізації запасів, підвищення ефективності маркетингових кампаній та покращення задоволеності клієнтів.

Дослідження підкреслило важливість інтеграції різних методів аналізу даних. Подальше дослідження може бути спрямоване на оптимізацію алгоритмів, розширення функціональності та адаптацію для різних типів кінотеатрів. Ця робота є значущим кроком у розумінні можливостей та впровадження методів аналізу даних у бізнесі кінотеатрів, сприяючи покращенню стратегічного управління та досягненню поставлених цілей.

4. Перелік джерел посилання

- [1] Big Data for Sustainable Development. URL: <https://www.un.org/en/global-issues/big-data-for-sustainable-development>
- [2] Часові ряди. URL: <https://kstat.pnu.edu.ua/wp-content/uploads/sites/63/2018/04/%D0%A7%D0%B0%D1%81%D0%BE%D0%B2%D1%96-%D1%80%D1%8F%D0%B4%D0%B8.pdf>
- [3] Кореляційний аналіз даних: Що це таке і як його застосовувати. URL: <https://ua5.org/model/1728-korelyaczijnj-analiz-danyh-shho-cze-take-i-yak-jogo-zastosovuvaty.html>
- [4] Які KPI промо-акції та як їх оцінити? URL: <https://datawiz.io/uk/blog/what-kris-of-promotion>
- [5] О. Є. Поморцева, Геоінформаційний аналіз, ХНУМГ ім. О. М. Бекетова, Харків, 2019.
- [6] А. В. Яковенко, О. О. Коновал, Управління IT-проектами, НТУУ «КПІ ім. І. Сікорського», Київ, 2017.
- [7] С. В. Мінухін, О. М. Беседовський, С. В. Знахур, Методи і моделі проектування на основі сучасних CASE-засобів, ХНЕУ, Харків, 2008.

Дослідження Методів та Технологій Автоматизації Процесів Формування Заявок для Веб-сайту Компанії з Прокату Авто

Діденко Денис^а, Віктор Левикін^а

^а Харківський національний університет радіоелектроніки, пр. Науки, 14, м. Харків, 61166, Україна

Анотація

Поява Інтернету принесла з собою багато корисних речей. Однією з них є можливість автоматизації оренди автомобіля. Завдання дослідити методи та технології автоматизації процесів формування заявок для сайту компанії з прокату автомобілів.

Ключові слова ¹

Прокат авто, автоматизація обліку заявок.

1. Вступ

За останні роки інтернет-бізнес в галузі прокату автомобілів став суттєвою складовою стратегії розвитку компаній, що надають послуги прокату транспортних засобів. Ця тенденція обумовлена зростанням попиту на мобільні та ефективні рішення для замовлення автомобілів через Інтернет. З цим виникає важлива проблема: як оптимізувати та автоматизувати процеси формування заявок на прокат авто для максимізації зручності клієнтів та підвищення конкурентоспроможності компаній у цьому динамічному ринковому середовищі. Сучасні клієнти вимагають не лише швидкості обслуговування, але й персоналізованого підходу та високого рівня сервісу. У цьому контексті, ефективність взаємодії з клієнтами та оптимізація процесів оформлення заявок на веб-сайтах стає стратегічною перевагою для компаній, які прагнуть відповідати сучасним стандартам якості обслуговування. Дослідження та впровадження нових методів та технологій автоматизації може відкрити нові можливості для удосконалення процесів формування заявок, забезпечуючи оптимальну реакцію на потреби клієнтів та забезпечуючи стабільний ріст бізнесу компаній, які діють у цьому секторі.

2. Аналіз існуючих методів та технологій

На сучасному етапі розвитку галузі прокату автомобілів існують різні підходи до формування заявок на веб-сайтах компаній, проте багато з них стикаються з визначеними обмеженнями. Традиційні методи, такі як заповнення форм на веб-сторінках чи застосунках, зберігаються у складних форматах, що часто може викликати незручності для користувачів та призводити до невірних введення інформації. Більшість існуючих систем покладаються на ручне введення даних клієнтами, що може вести до помилок та затримок у обробці заявок. Крім того, вони можуть бути обмеженими в функціональності, не надаючи достатньої гнучкості для врахування різноманітних потреб клієнтів. Деякі компанії вже застосовують деякі елементи автоматизації, такі як автоматичне заповнення деяких полів або використання систем рекомендацій на основі попередніх виборів клієнта. Проте, багато із цих рішень є фрагментарними та не використовують потенціал сучасних технологій. Дослідження зазначених обмежень та аналіз існуючих методів дозволить визначити прогалини в сучасних підходах до автоматизації формування заявок та покаже шляхи для подальшого розвитку та

Proceedings of the *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine
EMAIL: denys.didenko@nure.ua (A. 1); viktor.levykin@nure.ua (A. 2)
ORCID: 0009-0005-0613-4522 (A. 1); 0000-0002-7929-515X (A. 2)



вдосконалення цих процесів. Важливим етапом є також порівняння вартості, швидкості та точності різних методів для визначення оптимальних стратегій автоматизації.

3. Аналіз попередніх досліджень щодо автоматизації процесів взаємодії з клієнтами в галузі прокату автомобілів

Наукова спільнота вже виявила значущий інтерес до питань автоматизації процесів взаємодії з клієнтами в сфері прокату автомобілів. Декілька попередніх досліджень присвячені аспектам оптимізації та покращення відносин із споживачами в цьому конкретному контексті. Перш за все, підкреслюється важливість використання технологій для підвищення зручності клієнтів під час процесу прокату. Дослідження вказують на те, що ефективне використання мобільних додатків та веб-сайтів для оформлення заявок може значно поліпшити користувацький досвід та підвищити лояльність клієнтів. Однією з ключових тем є впровадження штучного інтелекту (ШІ) та машинного навчання (МН) у процеси автоматизації. Дослідники акцентують на можливостях аналізу даних для прогнозування попиту, персоналізації пропозицій та покращення стратегій ціноутворення. Також розглядається важливість інтеграції інтерактивних інтерфейсів та технологій, спрямованих на підвищення залучення користувачів. Дослідження показують, що інтерфейси, які сприяють взаємодії з клієнтами та надають рекомендації на основі їхніх індивідуальних потреб, можуть значно поліпшити якість обслуговування. Враховуючи ці попередні дослідження, стає зрозумілим, що переосмислення існуючих методів та використання новітніх технологій у сфері автоматизації процесів формування заявок може призвести до значного підвищення ефективності бізнес-процесів та забезпечити більш високий рівень задоволення клієнтів.

4. Висновки

У результаті аналізу існуючих методів та попередніх досліджень у галузі автоматизації процесів формування заявок для веб-сайтів компаній з прокату автомобілів можна зробити кілька важливих висновків.

По-перше, традиційні методи формування заявок мають свої обмеження, які включають незручності для користувачів та можливість виникнення помилок введення даних. Дослідження показують, що існуючі системи часто не використовують повністю потенціал сучасних технологій.

По-друге, важливим етапом є впровадження штучного інтелекту та машинного навчання для оптимізації процесів взаємодії з клієнтами. Автоматичний аналіз даних дозволяє підбирати індивідуальні пропозиції, передбачати попит та оптимізувати стратегії обслуговування.

По-третє, інтеграція інтерактивних інтерфейсів та технологій може значно покращити споживчий досвід. Взаємодія з клієнтами та надання персоналізованих рекомендацій сприяють підвищенню задоволеності користувачів.

З огляду на це, перспективи розвитку включають у себе впровадження інноваційних технологій, таких як розширений аналіз даних, а також постійне вдосконалення інтерфейсів для забезпечення зручності та доступності для широкого кола користувачів.

5. Література

- [1] Car rental. URL: https://en.wikipedia.org/wiki/Car_rental
- [2] Car renting. URL: <https://www.automotive-fleet.com/147063/car-renting-its-development-and-future>
- [3] Car Rental Insurance Tip Sheet. URL: https://www.consumeraffairs.com/travel/carrent_ins.html
- [4] History – The beginning of a moving tale. URL: <https://about.sixt.com/en/>
- [5] FluentValidation docs. URL: <https://docs.fluentvalidation.net/en/latest/>

Дослідження Інформаційної Технології Документообігу Для Оптимізації Роботи Ветеринарної Клініки

Надія Калита^a, Анастасія Соляник^a

^a Харківський національний університет радіоелектроніки, пр.Науки, 14, м.Харків, 61166, Україна

Анотація

Дослідження присвячено розгляду інформаційної технології документообігу у ветеринарних клініках та її використання для оптимізації робочих процесів та підвищення ефективності. Обговорюються сучасні тенденції у сфері документообігу, переваги та виклики застосування електронних систем, а також вплив цієї технології на якість медичного обслуговування у ветеринарних клініках.

Ключові слова ¹

автоматизація процесів, ветеринарна клініка, електронний документообіг, інформаційна технологія, клієнтське обслуговування, оптимізація робочих процесів, ефективність управління даними

1. Вступ

У сучасному інформаційному суспільстві, де технологічні інновації відкривають нові можливості для оптимізації різних сфер, важливим аспектом стає вдосконалення систем управління документообігом. Одним з напрямів, який потребує оптимізації у цьому контексті, є ветеринарна медицина. Актуальність дослідження цієї інформаційної технології обумовлена нагальною потребою в удосконаленні систем управління документообігом, зокрема у ветеринарних клініках. Ветеринарні клініки, як і інші заклади, стикаються з великим обсягом документації стосовно пацієнтів, персоналу, лікування, матеріально-технічного забезпечення, управління закладом тощо. Ефективна система електронного документообігу забезпечує всі переваги впровадження інформаційних технологій у процеси виробництва та адміністративного керування, що в свою чергу позитивно впливає на якість надання послуг.

2. Основна частина

Електронний документообіг – сукупність процесів, включаючи створення, обробку, відправлення, передавання, отримання, зберігання, використання та утилізацію електронних документів. Ці операції здійснюються з використанням перевірки цілісності та, за необхідності, з підтвердженням факту отримання таких документів [1].

Використання паперового документообігу, пов'язано з низкою проблем, а саме [2, 3]:

- витрати часу і ресурсів. Обробка паперових документів зазвичай вимагає значно більше часу порівняно з електронними системами. Ручне сортування та архівування паперових документів забирає велику кількість часу, що може впливати на продуктивність та ефективність робочих процесів;
- низька мобільність та доступність. Відсутність можливості швидкого доступу до паперових документів ускладнює роботу в умовах віддаленої роботи або у мобільному середовищі;

Information Systems and Technology (IST-2023), November 28 – December 1, 2023, Kharkiv, Ukraine

EMAIL: nadia.kalyta@nure.ua (A. 1); anastasiia.solianyk@nure.ua (A. 2)

ORCID: 0000-0001-6181-732X (A. 1); 0009-0007-8485-3623 (A. 2)



- втрати та пошкодження документів. Паперові документи піддаються ризику втрати або пошкодження в результаті пожежі, повені, крадіжки чи інших природних чи техногенних подій. Це може призвести до втрати важливої інформації та ускладнити відновлення даних;
- екологічні аспекти. Виробництво та використання паперу мають негативний вплив на довкілля;
- безпека і конфіденційність. Паперові документи можуть бути вразливі до несанкціонованого доступу та крадіжки, що створює загрозу конфіденційності інформації. На відміну від паперового, електронний документообіг передбачає захищені канали передавання інформації. Згідно з чинним законодавством електронні документи прирівнюються до оригіналів. На них накладаються електронні підписи (КЕП), які закон також прирівнює до власноручних підписів [4].

Метою роботи є дослідження електронного документообігу у ветеринарній клініці для визначення ефективних шляхів оптимізації робочих процесів. Для досягнення цієї мети необхідно провести комплексний аналіз стану електронного документообігу у ветеринарній клініці. Початковим етапом є визначення передумов для впровадження ефективної системи, зокрема, врахування технологічних можливостей та особливостей клінічної практики. Проведення детального аналізу функцій кожної підсистеми інформаційної системи (ІС) документообігу допоможе визначити їхню роль і місце в загальній структурі. Необхідно визначити потреби користувачів, розглянути індивідуальні особливості робочих процесів кожної категорії працівників клініки. Це дозволить налаштувати систему так, щоб вона відповідала конкретним вимогам кожного користувача.

На сьогоднішній день вже існують деякі рішення електронного документообігу у ветеринарних клініках. Наприклад, Hippo Manager – програмне забезпечення для керування ветеринарною клінікою, яке просте у використанні та призначено для підвищення ефективності, фіксації витрат клініки і отримання доступу до даних у реальному часі [5]. Ця система включає в себе головні можливості електронного документообігу, такі як електронна медична історія, планування прийому, облік клієнтів та їхніх тварин, електронні рецепти, а також:

- вбудовані звіти. Дозволяють аналізувати важливі показники ефективності роботи закладу [6];
- контроль запасів. Дозволяє підтримувати рівень запасів, точно відстежувати прибуток, надаючи всі інструменти для керування запасами [6];
- портал пацієнтів. Клієнти ветеринарної клініки можуть отримати доступ до медичних записів тварин онлайн, переглядати рахунки, друкувати сертифікати вдома [6];
- таблиці лікування. Автоматизовані листи, які спрощують переміщення даних про тварин та відстежують процес та прогрес їх лікування у клініці [6].

eVetPractice.com – це повнофункціональне ветеринарне програмне забезпечення, призначене для обслуговування стартапів та підприємств. eVetPractice.com надає комплексні рішення, розроблені для веб-застосунків. Ця онлайн ветеринарна система пропонує плани лікування, планування відвідувань, управління запасами, управління рецептами, виставлення рахунків [7].

Інформаційна система документообігу включає ряд підсистем, спрямованих на управління медичною документацією та забезпечення ефективного функціонування клінічного процесу, кожна з яких відповідає за конкретні функції:

- підсистема медичних карток пацієнтів. Функції: зберігання та оновлення медичних записів кожного пацієнта, внесення даних про діагнози, лікування та профілактичні заходи, оформлення направлень;
- підсистема замовлення та управління медикаментами. Функції: оформлення рецептів та замовлення ліків, відстеження залишків ліків та управління ними;
- підсистема запису на прийом. Функції: запис пацієнтів на прийом до ветеринара, відстеження графіка роботи ветеринарів, підтвердження та скасування прийомів;
- підсистема лабораторних досліджень. Функції: замовлення та відстеження результатів лабораторних аналізів, зберігання лабораторних даних у пацієнтських картках;
- підсистема рахунків та фінансів. Функції: виставлення рахунків за надані послуги, облік фінансових транзакцій та оплат;
- підсистема зберігання та керування документами. Функції: зберігання та організація всіх видів документів клініки, доступ до стандартів та регуляторних актів.

Ці підсистеми утворюють ІС документообігу ветеринарної клініки, спрямовану на покращення управління медичною документацією, підвищення ефективності роботи та забезпечення високого стандарту надання ветеринарних послуг.

Користувачі ІС ветеринарної клініки можуть включати адміністраторів клініки, лікарів, медсестер, клієнтів (власників тварин). Кожна категорія користувачів має визначені ролі та рівень доступу відповідно до їхніх функціональних обов'язків:

- адміністратори клініки. Функції: реєстрація клієнтів, управління акаунтами власників та тварин, фінансові операції, запис клієнтів на прийоми, підтвердження та скасування записів. Доступ до даних клієнтів та їхніх тварин, можливість вносити до них зміни, доступ до графіків прийому та внесення до них змін;
- ветеринарні лікарі. Функції: обробка медичних записів та призначення лікування. Доступ до електронних медичних карт тварин, можливість вносити зміни у медичних записах;
- медсестри. Функції: запис та оновлення певних медичних даних тварин. Обмежений доступ до медичних записів;
- власники тварин. Функції: доступ до особистої інформації про своїх тварин, включаючи медичні записи, візити, рецепти, запис на візити та перегляд рекомендацій лікаря. Доступ до даних тільки щодо їх тварин, з обмеженими правами на редагування.

Ці ролі та рівні доступу допомагають забезпечити безпеку даних та ефективну обробку інформації у ветеринарній клініці, відповідно до функціональних обов'язків кожного типу користувача.

Впровадження електронного документообігу у ветеринарних клініках може стикатися з деякими викликами, такими як необхідність адаптації персоналу до нових технологій, труднощі інтеграції з існуючими системами, витрати часу на перехід від паперового до електронного формату, створення стандартів та забезпечення сумісності між різними системами, а також потреба в надійності та постійному доступі до системи. Оновлення та підтримка електронного документообігу також можуть вимагати додаткових ресурсів та уваги для забезпечення ефективності та актуальності системи. Забезпечення ефективного вирішення цих викликів є критичним для успішної імплементації електронного документообігу. Воно повинно бути спрямоване на максимізацію переваг, що може допомогти ветеринарним клінікам оптимізувати робочі процеси, підвищити точність та доступність медичної інформації, що призведе до покращення якості обслуговування та задоволення клієнтів.

3. References

- [1] Про електронні документи та електронний документообіг : Закон України від 22.05.2003 р. № 851-IV : станом на 1 серп. 2022 р. URL: <https://zakon.rada.gov.ua/laws/show/851-15#Text>
- [2] Паперовий VS електронний документообіг. URL: <https://vesti.dp.ua/papеровий-vs-elektronnij-dokumentooBIG>
- [3] 5 причин, чому паперовий документообіг у минулому. URL: <https://medoc.ua/blog/5-prichin-chomu-papеровий-dokumentooBIG-u-minulomu>
- [4] Що таке електронний документообіг. Вчасно. URL: <https://vchasno.ua/shcho-take-elektronnyi-dokumentooBIG>
- [5] Veterinary Practice Management Software. Hippo Manager. URL: <https://www.hippomanager.com>
- [6] Best In Class Veterinary Software Features. Hippo Manager. URL: <https://www.hippomanager.com/solutions>
- [7] eVetPractice.com Pricing, Features, and Reviews (Dec 2023). URL: <https://www.softwaresuggest.com/evetpractice-com>

Застосування штучного інтелекту в мультимедійних системах

Ірина Кириченко^а та Дмитро Заїграєв^а

¹ Харківський національний університет радіоелектроніки, проспект Науки, 14, Харків, Харківська область, Україна, 61166

Анотація

Стаття досліджує вплив та застосування методів ШІ в мультимедійних системах, охоплюючи аналіз відображення, обробки та розуміння мультимедійного контенту. Вона також розглядає технології, що використовують ШІ для покращення сприйняття та взаємодії користувача з мультимедійними ресурсами.

Ключові слова ¹

Штучний інтелект, мультимедійні системи, обробка мультимедіа, відображення контенту, розпізнавання образів, аналіз звуку, мультимедійна взаємодія

1. Вступ

Мультимедійні системи змінили спосіб, яким ми сприймаємо інформацію, розважаємося та взаємодіємо. Вони включають у себе різноманітні типи контенту, такі як зображення, відео, аудіо та текстові дані, які співіснують, створюючи захоплюючі враження. Розвиток штучного інтелекту (ШІ) суттєво вплинув на можливості та потенціал мультимедійних систем. Інтеграція методів ШІ у мультимедійні системи відкрила безліч можливостей, дозволяючи цим системам не лише обробляти, а й розуміти, тлумачити та покращувати представлений контент. Ця взаємодія між ШІ та мультимедіа породжує інновації, що виходять за рамки традиційних обмежень, збагачуючи користувацькі враження та розширюючи можливості творчого виразу.

Штучний інтелект у мультимедійних системах виявився ключовим фактором, що дозволяє вдосконалювати та розширювати функціональність системи. Відповідно до цього, алгоритми машинного навчання та нейронні мережі використовуються для розпізнавання образів, розуміння мови, аналізу звуку та створення інтелектуальних систем, що автоматично адаптуються до потреб користувачів. Інтелектуальні системи в мультимедіа розширюють можливості не лише для споживачів, але й для творців контенту. За допомогою аналізу даних про користувачів, ці системи можуть прогнозувати та створювати персоналізований контент, який краще відповідає індивідуальним уподобанням аудиторії.

Штучний інтелект став ключовим каталізатором для інновацій у сфері мультимедіа, змінюючи спосіб, яким ми сприймаємо, створюємо та споживаємо контент. В цій статті ми докладніше розглянемо вплив та перспективи розвитку ШІ у мультимедійних системах, розкриваючи, як ці технології моделюють сучасні реалії та перетворюють майбутність цієї сфери.

2. Розпізнавання образів за допомогою ШІ

Останні дослідження в галузі розпізнавання образів фокусуються на вдосконаленні нейронних мереж та моделей глибокого навчання. Ключові досягнення полягають у використанні архітектур CNN та моделей, таких як ResNet, EfficientNet, які відзначаються високою точністю

Information Systems and Technology (IST-2023), November 28 – December 1, 2023, Kharkiv, Ukraine
EMAIL: iryna.kyrychenko@nure.ua (A. 1); dmytro.zaihraiev@nure.ua (A. 2)
ORCID: 0000-0002-7686-6439 (A. 1)



класифікації та розпізнавання об'єктів на зображеннях. Недавні дослідження також зосереджуються на розвитку моделей, які забезпечують відновлення зображень високої якості з використанням генеративних адверсарних мереж (GANs) та анотування зображень для поліпшення точності та різноманітності даних.

Покращення якості обробки та редакція зображень. Сучасні підходи до обробки зображень використовують глибокі нейронні мережі для автоматичної корекції зображень. Вони можуть не лише підвищувати роздільну здатність, але й виправляти дефекти, такі як розмиття, розтягнення та забарвлення [5]. Деякі методи також активно використовують структури зворотного зв'язку для покращення якості зображень. Особливий акцент зроблений на розвиток автоматизованих методів реставрації та усунення шуму, що забезпечують високоякісне зображення після обробки.

Персоналізована обробка контенту. Штучний інтелект дозволяє створювати індивідуальні налаштування обробки зображень, враховуючи унікальні уподобання користувачів. Системи можуть адаптуватися до вимог щодо кольору, контрастності, яскравості та стилю зображення, забезпечуючи користувачам персоналізований досвід [4].

Останні розробки у цій області включають в себе розвиток алгоритмів, що враховують контекст використання зображень, враховуючи не лише особливості користувача, але й особливості використання зображення в різних сценаріях.

3. Універсальність застосування

Технології розпізнавання образів на основі ШІ мають великий потенціал у різних галузях. Вони використовуються для розваг, медицини, промисловості, безпеки та багатьох інших сфер. Сучасні дослідження активно досліджують можливості застосування ШІ в архітектурі "інтернету речей" (IoT), що дозволяє автоматизувати процеси та роботу різних пристроїв, використовуючи візуальний аналіз та розпізнавання образів [3].

1. Розваги та медіа - штучний інтелект дозволяє створювати індивідуальні налаштування обробки зображень, враховуючи унікальні уподобання користувачів. Системи можуть адаптуватися до вимог щодо кольору, контрастності, яскравості та стилю зображення, забезпечуючи користувачам персоналізований досвід [2]. Останні розробки у цій області включають в себе розвиток алгоритмів, що враховують контекст використання зображень, враховуючи не лише особливості користувача, але й особливості використання зображення в різних сценаріях.

2. Медицина та діагностика - в медицині технології розпізнавання образів використовуються для діагностики захворювань на основі медичних зображень. Моделі глибокого навчання допомагають ідентифікувати патології на знімках рентгенівських та комп'ютерної томографії, що дозволяє швидше та точніше поставити діагнози.

3. Промисловість та автоматизація - у промисловості технології розпізнавання образів використовуються для візуального контролю якості продукції на виробничих лініях [1]. Це дозволяє виявляти дефекти або невідповідності стандартам та автоматично вилучати неякісні вироби з лінії виробництва.

4. "Інтернет речей" та візуальний аналіз - у сфері "інтернету речей" системи розпізнавання образів застосовуються для розуміння оточуючого середовища. Датчики та камери здатні аналізувати та розпізнавати об'єкти, що допомагає в автоматизації процесів, наприклад, у системах "розумного будинку" або в автономних транспортних засобах.

Таблиця 1

Приклади застосування технологій розпізнавання образів у різних галузях

Галузь	Приклад застосування
Розваги та медіа	Віртуальна реальність та розширена реальність (VR/AR)
Медицина та діагностика	Рентгенологія та комп'ютерна томографія (СТ)
Промисловість	Візуальний контроль якості виробництва
"Інтернет речей"	"Розумний дім"

Приклади застосування технологій розпізнавання образів у різних галузях наведені у Табл.1. Розваги та медіа - віртуальна реальність (VR) та розширена реальність (AR). Технології розпізнавання образів використовуються для створення вражаючих візуальних ефектів у віртуальній реальності. Наприклад, додатки, що використовують AR для розпізнавання об'єктів у реальному світі та проєктують на них інтерактивну інформацію.

Медицина та діагностика - рентгенологія та комп'ютерна томографія (СТ). Технології розпізнавання образів у медицині допомагають лікарям швидше та точніше виявляти захворювання. Алгоритми використовуються для автоматичного розпізнавання патологій на знімках рентгеновських або СТ зображень, що полегшує діагностику.

Промисловість та автоматизація - візуальний контроль якості виробництва. У промисловості системи розпізнавання образів використовуються для візуального контролю якості виробництва. Камери та алгоритми сприймають та аналізують зображення в реальному часі, виявляючи дефекти або нестандартність у виробках.

"Інтернет речей" та візуальний аналіз – «розумний будинок». У системах "розумного будинку" камери та датчики використовуються для розпізнавання образів та аналізу оточуючого середовища. Наприклад, системи визначення обличчя людей для автоматичного включення/вимикання підсвітки, контролю доступу та підтримки безпеки.

4. Висновки

Таким чином, вплив штучного інтелекту на розвиток мультимедійних систем, зокрема у сфері розпізнавання образів. ШІ відкриває безліч можливостей для покращення та розширення функціональності мультимедійних систем. Інтеграція методів ШІ дозволяє системам не лише обробляти, а й розуміти, тлумачити та покращувати представлений контент, що породжує інновації та збагачує користувацькі враження. Розпізнавання образів через ШІ вказують на значний прогрес у цій області. Використання нейронних мереж та глибокого навчання сприяє високій точності класифікації та розпізнавання об'єктів на зображеннях. Крім того, застосування генеративних адверсарних мереж та алгоритмів автоматичної корекції зображень підвищує якість обробки та дозволяє створювати персоналізований контент. Штучний інтелект не лише впливає на розваги та медіа, де він дозволяє створювати індивідуальний досвід для користувачів, але також знаходить застосування у сферах медицини, промисловості та "інтернету речей". В медицині він сприяє точнішій діагностиці, у промисловості - контролю якості, а в "інтернеті речей" - автоматизації та розумінню оточуючого середовища.

Загалом, штучний інтелект у мультимедійних системах відкриває широкі горизонти для творчого виразу, забезпечуючи нові можливості як для споживачів, так і для творців контенту, та стає ключовим фактором, що моделює сучасні реалії та перетворює майбутність цієї сфери.

5. Список джерел

- [1] Kyrychenko, I., Proniuk, G., Geseleva, N., & Tereshchenko, G. (2019). "3rd International Conference on Computational Linguistics and Intelligent Systems (COLINS-2019)." Kharkiv, Ukraine, April 18-19, 2019. Volume 2362, pp. 266-276. ISSN 16130073.
- [2] Kyrychenko, I., & Roshka, V. (2022). "Proceedings of the 6th International Scientific and Practical Conference SCIENTIFIC COMMUNITY: INTERDISCIPLINARY RESEARCH." Hamburg, Germany, 26-28 January 2022, pp. 1065-1069.
- [3] Wang, T. C., Liu, M. Y., Zhu, J. Y., Tao, A., Kautz, J., & Catanzaro, B. (2018). High-resolution image synthesis and semantic manipulation with conditional GANs. In Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), pp. 8798-8807.
- [4] Zhang, K., Van Der Maaten, L., & Chen, Z. (2019). Self-tuning networks: Bilevel optimization of hyperparameters using structured best-response functions. In Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), pp. 10513-10522.
- [5] Zhang, R., Isola, P., & Efros, A. A. (2018). Colorful image colorization. In European Conference on Computer Vision, pp. 649-666. Springer, Cham.

Дослідження сучасних технологій оптимізації вебзастосунків

Єгор Лебідь^a, Юрій Міщераков^a

^a Харківський національний університет радіоелектроніки, пр. Науки 14, Харків, 61166, Україна

Анотація

У рамках роботи проведено аналіз критеріїв оптимізації вебзастосунків та визначено актуальні критерії для виміру продуктивності системи. Створено опис математичної моделі задачі у вигляді мінімізації та максимізації відповідних критеріїв. Проведено аналіз і визначено переваги, недоліки та актуальність застосування кожного з підходів до масштабування.

Ключові слова ¹

Критерії оптимізації, масштабування, вебзастосунки, інформаційні системи.

1. Вступ

Наразі інформаційні системи є надзвичайно популярними. Вони широко використовуються у різних сферах та галузях, оскільки вирішують ряд проблем завдяки своїм перевагам, таким як автоматизація процесів та зберігання великих обсягів даних. Вебзастосунки – це програми, які працюють на веб-серверах і доступні користувачам через веббраузери. Вони призначені для надання користувачам зручного інтерфейсу для взаємодії з сервером і виконання різних завдань. На сьогоднішній день вебзастосунки використовуються кожною людиною, яка має доступ до мережі. Через глобальне використання мережі Інтернет багато компаній зацікавлена у створенні власного вебзастосунку для надання тих чи інших послуг або продажу товарів.

Актуальність теми полягає у тому, що оптимізація вебзастосунків відіграє ключову роль для конкурентоспроможності та забезпечення комфортного досвіду користування.

2. Результати дослідження критеріїв оптимізації та визначення узагальненого критерія оптимізації

В результаті аналізу було визначено перелік ключових критеріїв оптимізації вебзастосунків:

1. пропускна здатність (throughput);
2. доступність (availability);
3. надійність (reliability);
4. затримка (latency).

Критерій пропускної здатності (throughput) характеризує кількість транзакцій або запитів, які може обробити вебзастосунок за одиницю часу. Висока пропускна здатність, як правило, свідчить про кращу продуктивність та масштабованість[1]. Формула для розрахунку пропускної здатності в продуктивності веб-додатків має такий вигляд:

$$RPS = \frac{N_R}{T}, \quad (1)$$

де RPS – пропускна здатність запитів, кількість запитів за секунду; N_R – кількість запитів за проміжок часу T .

Proceedings of the *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine

EMAIL: yehor.lebid@nure.ua (A. 1); iurii.mishcheriakov@nure.ua (A. 2)

ORCID: 0009-0001-2552-6696 (A. 1); 0000-0002-5334-1808 (A. 2)



Критерій доступності (availability) – це ключовий показник, що характеризує стабільність та працездатність системи у нестабільні періоди. Таким чином, високий показник доступності, як правило, свідчить про кращу продуктивність, масштабованість та стабільність роботи[2]. Формула для розрахунку має наступний вигляд:

$$A = \frac{T_U}{T_U + T_D} * 100, \quad (2)$$

де A – критерій доступності у відсотках; T_U – час роботи застосунку (uptime); T_D – час, коли застосунок недоступний (downtime).

Критерій надійності (reliability) характеризує здатність системи стабільно працювати відповідно до своїх специфікацій. Це ймовірність безвідмовної роботи програмного забезпечення протягом певного періоду часу у визначеному середовищі. Одним із поширених способів вираження надійності є середній час напрацювання на відмову (MTBF):

$$MTBF = \frac{T}{N}, \quad (3)$$

де $MTBF$ – середній час напрацювання на відмову; T – загальний час роботи системи; N – кількість відмов.

Критерій затримки (latency) в контексті продуктивності вебзастосунків характеризує затримку або час, необхідний для передачі даних між пристроєм користувача і вебсервером. Це час, необхідний для передачі даних з однієї точки мережі в іншу[3]. Формула для розрахунку має наступний вигляд:

$$L = T_t - T_p, \quad (4)$$

де L – латентність, відноситься до часу затримки між ініціюванням запиту або операції та отриманням відповіді або результату; T_t – час, необхідний для виконання запиту або операції, з моменту його ініціювання до моменту отримання відповіді; T_p – час, витрачений на фактичну обробку даних або обчислення, без урахування часу очікування або простою.

Провівши аналіз критеріїв оптимізації вебзастосунків та визначивши перелік основних критеріїв, що будуть використані, можна сформулювати основну математичну модель задачі оптимізації. Метою оптимізації є зменшення одних показників оптимізації і підвищення інших. Так в рамках даної моделі критеріями, що вимагають мінімізації є:

- затримка (latency), оскільки час очікування треба зменшити для кращого користувацького досвіду, критерій прямує до мінімуму: $L \rightarrow \min$.

Серед критеріїв, що вимагають максимізації це:

- пропускна здатність (throughput), оскільки обробка більшої кількості запитів за одиницю часу має гарантувати високу працездатність застосунку під час надмірного навантаження, критерій прямує до максимуму: $RPS \rightarrow \max$;
- доступність (availability), оскільки час доступності застосунку для використання підвищує якість користувацького досвіду та репутацію компанії в цілому, критерій прямує до максимуму: $A \rightarrow \max$;
- надійність (reliability), оскільки надійність системи забезпечує безперебійну роботу системи під час незапланованих проблем із системою, критерій прямує до максимуму: $MTBF \rightarrow \max$.

Отже, визначивши критерії, що вимагають мінімізації та максимізації, можна представити узагальнений критерій оптимізації у наступному вигляді:

$$O = \{ \min(L), \max(RPS), \max(A), \max(MTBF) \}, \quad (5)$$

де L – критерій латентності (latency); RPS – критерій пропускної здатності (throughput); A – критерій доступності (availability); $MTBF$ – критерій доступності (availability), що характеризує середній час напрацювання на відмову.

3. Результати дослідження підходів до масштабування

Одним з найбільш популярних та сучасних методів оптимізації вебзастосунків є технологія масштабування. Існує два види масштабування, серед яких вертикальне та горизонтальне.

Вертикальне масштабування – це процес збільшення потужності існуючої системи, як наприклад процесора або оперативної пам'яті, для задоволення зростаючих вимог до системи. Це – процес покращення одного серверу. Даний підхід до масштабування є досить простим, оскільки не вимагає жодних змін кодової бази, а просто вимагає залучення потужнішого ресурсу для виконання. Для певних потреб даного підходу може бути достатньо, але існують деякі обмеження такого підходу. І головне обмеження полягає в тому, що не існує фізичної можливості забезпечити постійне вертикальне масштабування системи.

Для детального опису даного обмеження необхідно розглянути закон Мура, що був названий Гордоном Муром, одним із співзасновників компанії «Intel», яка спеціалізується на розробці процесорів для комп'ютерів. Закон полягає в тому, що кількість транзисторів в інтегральній схемі подвоюється кожні два роки, а вартість зменшується в два рази. Для спрощення розгляду кількість транзисторів будуть характеризувати швидкість виконання алгоритмів. Отже, ми можемо зробити припущення, базуючись на даному законі, що швидкість процесорів буде збільшуватись у два рази кожні два роки. Це означає що ріст швидкості має бути експоненціальний, але у сучасному світі ми бачимо, що графік росту швидкості процесорів підходить до свого плато. За останні 10 років ріст швидкості значно скоротився[4]. Схематично це подано на рисунку 1.

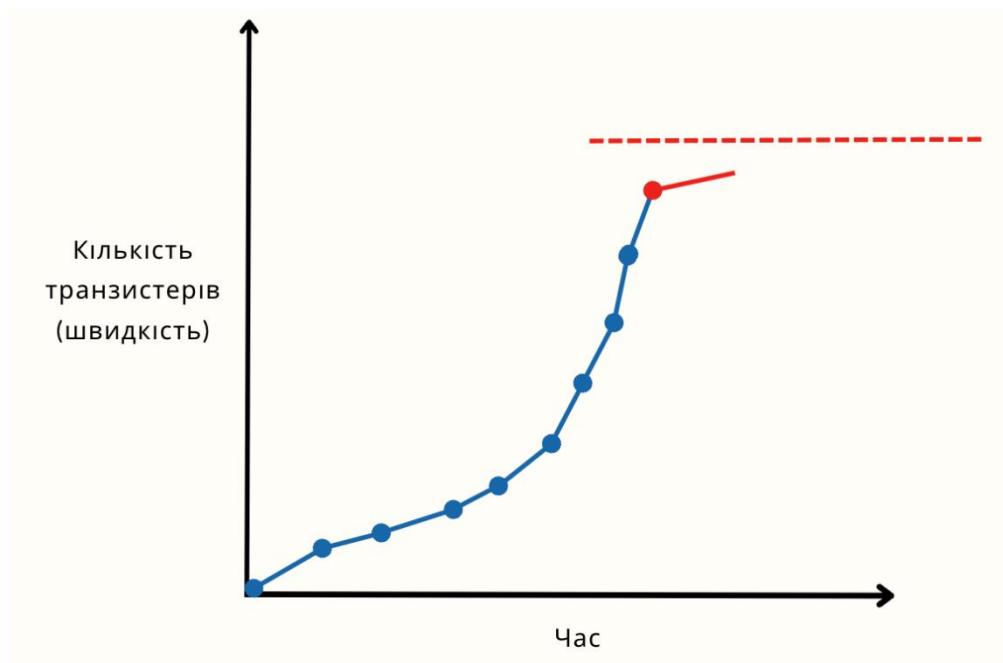


Рисунок 1: Схематичне представлення закону Мура

Враховуючи дане обмеження, можна зробити висновок, що вертикальне масштабування має дане обмеження у якісному покращенні одного сервера, оскільки так чи інакше темпи росту швидкості роботи процесору значно зменшуються.

Іншим ключовим підходом до масштабування є горизонтальне масштабування. Воно являє собою підхід до розширення або підвищення продуктивності системи шляхом додавання нових ресурсів, зазвичай збільшення кількості обчислювальних вузлів. Це відбувається шляхом збільшення кількості серверів, які використовуються для роботи системи[5]. Цей підхід є більш ефективним у рамках масштабування, оскільки практично не має обмежень. Окрім того даний підхід покращує й інші критерії оптимізації вебзастосунків, такі як пропускна здатність (throughput), доступність (availability) та надійність (reliability) завдяки більшій кількості вузлів в системі. По-друге, даний підхід вимагає особливої уваги до побудови архітектури вебзастосунку, оскільки вимагає підтримки розгалуженої обробки інформації в застосунку.

Враховуючи усі переваги і недоліки[6] в основних характеристиках вертикального і горизонтального масштабування було виконано порівняння, що надане у таблиці 1.

Таблиця 1

Характеристики вертикального та горизонтального масштабування системи

Характеристика	Вертикальне масштабування	Горизонтальне масштабування
Розподіл ресурсів	Один сервер	Кілька серверів
Масштабованість	Обмежена	Практично необмежена
Доступність вебзастосунку під час оновлення	Ні, оскільки має один сервер	Так, оскільки сервери оновлюються по черзі
Стійкість до відмов	Низька через єдину точку відмови	Висока через кількість серверів
Складність в реалізації	Низька	Висока
Продуктивність	Підвищує продуктивність одного сервера.	Розподіляє навантаження між серверами
Пропускна здатність	Незначно підвищується	Значно підвищується
Доступність	Незначно підвищується	Значно підвищується
Надійність	Незначно підвищується	Значно підвищується
Затримка	Незначно зменшується	Значно зменшується

4. Висновки

В результаті дослідження методів вирішення задачі було проведено аналіз критеріїв оптимізації вебзастосунків та визначено актуальні критерії для виміру продуктивності системи. Після дослідження метрик було створено опис математичної моделі задачі у вигляді мінімізації та максимізації відповідних критеріїв та описано узагальнений критерій оптимізації. Було проведено аналіз і визначено переваги та недоліки кожного з підходів до масштабування вебзастосунків. Були зроблені висновки щодо актуальності застосування кожного з них.

5. Література

- [1] T. Guild, What is Throughput in Performance Testing, 2022. URL: <https://testguild.com/performance-testing-what-is-throughput>.
- [2] Medium, The Five Key Concepts of Software Architecture: Scalability, Availability, Reliability, Performance, and Resilience, 2023. URL: <https://medium.com/@fullstacktips/the-five-key-concepts-of-software-architecture-scalability-availability-reliability-8927e7d0c6fe>.
- [3] Cloudflare, What is latency, 2023, URL: <https://www.cloudflare.com/en-gb/learning/performance/glossary/what-is-latency>.
- [4] M. Roser, What is Moore's Law, 2023. URL: <https://ourworldindata.org/moores-law>.
- [5] M. Kleppmann, Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems 1st ed., Gravenstein Highway North, Sebastopol, CA 2017.
- [6] B. Burns, Designing Distributed Systems: Patterns and Paradigms for Scalable, Reliable Services 1st ed., Gravenstein Highway North, Sebastopol, CA 2018.

Дослідження Текстових Генеративних Систем з Асоціативною Пам'яттю

Євгеній Мамочка^a, Андрій Єрохін^a

^a Харківський Національний Університет Радіоелектроніки, 14, просп. Науки., Харків, 61166, Україна

Анотація

Дослідження присвячене підвищенню ефективності роботи текстових генеративних систем. Проведено аналіз мовних систем з різними типами пам'яті та виявлені особливості використання асоціативної пам'яті в генеративних системах. В роботі розглядаються: система без архітектурних змін та система зі змінним типом пам'яті, яка крім основної моделі GPT-2 має додаткові шари.

Ключові слова

Асоціативна пам'ять, пам'ять, нейронні мережі, магістерське дослідження

1. Вступ

На даний момент існує багато доступних систем генерації тексту, які відрізняються за якістю в залежності від моделі. Якість тексту прямо пропорційна якості моделі, яка, у свою чергу, залежить від розміру, типу та обсягу навчальних даних, а також типу пам'яті системи.

Теоретично асоціативна пам'ять у мовних моделях передбачає навчання пам'яті цих моделей без вчителя, що спрощує генерацію тексту в цілому, а також розробку таких систем. Це дослідження спрямоване на підвищення ефективності генерації тексту з використанням асоціативної пам'яті.

2. Аналіз можливостей використання асоціативної пам'яті в текстових генеративних системах

Нейронна асоціативна пам'ять (НАП) - це нейромережеві моделі, що складаються з нейроноподібних та синапсоподібних елементів. У будь-який момент часу стан нейронної мережі задається вектором нейронної активності, який називається паттерном активності. Нейрони оновлюють значення своєї активності на основі вхідних даних, які вони отримують (через синапси). У найпростіших моделях нейронних мереж функція входу-виходу нейрона є функцією ідентичності або порогова операція. Треба зауважити, що в останньому випадку стан нейронної активності є бінарним: активний або неактивний. Інформація, яку обробляє нейронна мережа, представляється у вигляді шаблонів активності - паттернами активності (наприклад, представлення дерева може бути паттерном активності, де активні нейрони відображають зображення дерева). Таким чином, шаблони активності є представленнями елементів, що обробляються у мережі. Представлення називається розрідженим, якщо співвідношення між активними та неактивними нейронами невелике.

Синапси в нейронній мережі - це зв'язки між нейронами або між нейронами та волокнами, що несуть зовнішній вхід.

Нейрон - це одиниця, що виконує обчислення. Вона отримує дані з вхідного шару, виконуючи з нею прості обчислення, а потім передаючи наступному нейрону.

Proceedings of the *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine

EMAIL: evhenii.mamochka@nure.ua (A. 1); andriy.yerokhin@nure.ua (A. 2)

ORCID: 0009-0005-3358-4553 (A. 1); 0000-0002-8867-993X (A. 2)



Information Systems and Technology (IST-2023) Part 2. Youth section

Різні функції пам'яті визначаються тим, як можна отримати доступ до вивчених шаблонів вибірково звертаючись до вхідного зразка. Функція розпізнавання зразків означає класифікування вхідних зразків на два класи: знайомі зразки та решту. Асоціація шаблонів описує функцію асоціювання певних вхідних зразків з певними зразками у пам'яті, тобто кожна пам'ять складається з пари вхідного і бажаного вихідного зразків.

Асоціативна пам'ять у нейронних мережах стосується їхньої здатності запам'ятовувати та асоціювати вхідні дані чи образи. Цей термін охоплює кілька аспектів:

1) пам'ять образів – мережі можуть запам'ятовувати та розпізнавати конкретні образи у вхідних даних, виявляючи шаблони та закономірності;

2) асоціативне запам'ятовування – мережі відновлюють інформацію з пам'яті, користуючись частинами вхідних даних або ключовими елементами, що дозволяє їм асоціювати вхідні дані з певними виходами;

3) навчання з учителем і без учителя - мережі можуть навчатися з учителем, коли надаються пари вхід-вихід, або без учителя, коли вони самостійно виявляють закономірності в даних;

4) зіставлення схожих образів – система порівнює вхідні дані з відомими зразками, що дозволяє їй робити висновки або прогнози на основі схожості.

Перевага нейронної асоціативної пам'яті над іншими алгоритмами зберігання шаблонів, наприклад, таблицями пошуку хеш-кодів, полягає в тому, що доступ до пам'яті може бути відмовостійким по відношенню до змін вхідного шаблону. Для асоціації шаблонів це означає, що вихідний шаблон може бути створений для набору вхідних шаблонів, які є найближчі до вхідного шаблону, представленого під час навчання.

Нейронні мережі можуть досить ефективно працювати з різними типами даних, такими як зображення, звук, відео, текст тощо. Тоді з точки зору підвищення ефективності роботи текстових генеративних систем важливою та актуальною задачею є саме дослідження асоціативного типу пам'яті. Використання асоціативної пам'яті у мовних моделях є доцільним, оскільки це спрощує перевірку асоціацій між словами та контекстом під час генерації тексту.

3. Мета дослідження і постановка задачі дослідження

Головна мета дослідження – порівняння двох мовних систем з різними типами пам'яті та виявлення доцільності використання саме асоціативної пам'яті в генеративних системах.

Задачі, які необхідно виконати під час проведення дослідження:

- 1) знайти мовну модель, яка відповідає конкретним вимогам дослідження;
- 2) адаптувати обрану мовну модель для роботи з асоціативною пам'яттю, що становить основу дослідження;
- 3) детально задокументувати процес розробки системи;
- 4) підготувати дані для навчання мовної системи;
- 5) провести навчання розроблених систем на конкретних вхідних даних;
- 6) порівняти дві отримані нейронні мережі на однакових даних з метою визначення ефективності використання асоціативної пам'яті.

Навчання даних систем виконувалось на сервері Microsoft Azure.

Через велику кількість даних в матрицях, було визнано доцільним розпаралелити дії, для чого було використано один із видів оптимізації, а саме – TPU.

Для перевірки даних систем використовувалась метрика *accuracy*, а саме – порівняння очікуваних даних і даних нейронної мережі.

4. Реалізація задач дослідження

Отримана система є трансформером GPT-2, тобто системою для обробки текстових послідовностей, таких як у завданнях машинного перекладу та автоматичного реферування. Основна його особливість – механізм уваги, який виявляє взаємозв'язки між частинами вхідних і вихідних даних. Модель GPT-2 використовує *transformer attention*, що дає змогу ефективно обробляти вхідні дані, враховуючи всі позиції в послідовності.

У результаті ми отримали дві системи: одна система без архітектурних змін, тобто має свій звичайний вид GPT-2, а друга система – зі зміненим типом пам'яті, тобто окрім основної моделі GPT-2 має додаткову мультимодальну "голову" входу. Вдосконалена система має такі додаткові шари (1):

Memory output -> Embedding -> Bidirectional LSTM -> Attention Mechanism. (1)

де *Memory output* – вивід додаткової моделі без вчителя, яка повинна створювати асоціації на основі автосоціативного принципу і поновлювати образи даних.

Спочатку дані обробляються системою пам'яті, далі дані стискаються за подібним змістом. На наступному кроці двонаправлена LSTM виконує роботу з даними (прогнозування послідовності, класифікація даних, розпізнавання іменованих даних, позначення частин мови), а на останньому кроці – вибираються важливі данні.

У результаті до двох розроблених систем ставляться однакові запитання (а також N - грами в процесі навчання). У трансформерів є ймовірність наступного слова: якщо текст буде написаний з більшим значенням ймовірності, то він вважається більш правильним також, якщо метрика точності буде високою. Тобто дві моделі дають схожу відповідь, але там, де точність відповіді більше, то там і краще текст. Окрім того враховується її відповідність до даних, на яких система навчалася.

5. Висновки

В результаті проведення дослідження виявилось, що асоціативна пам'ять є особливо ефективною при роботі з завданнями, які передбачають роботу з контекстом, інтерпретацію великих обсягів інформації та генерацію змісту з урахуванням попередніх знань. Проте ефективність цього методу може варіюватися в залежності від конкретної задачі, і деякі завдання можуть бути менш чутливими до використання асоціативної пам'яті.

6. Література

- [1] Haykin S. Neural Networks. A comprehensive Foundation. - 2ed. - 1999. -P. 690.
- [2] Zhao H. Global asymptotic stability of Hopfield neural network involving distributed delays - 2004. - Vol. 17, N 1. - P. 48 - 53.
- [3] Rini,J.M., Schulze-Gahmen,U. and Wilson,I.A. (1992) Science, 255, 959–965.
- [4] Narayanan,A., Sellers,B.D. and Jacobson,M.P. (2009) J. Mol. Biol., 388, 941–953.
- [5] Martin,A.C., Cheetham,J.C. and Rees,A.R. (1989) Proc. Natl Acad. Sci. USA, 86, 9268–9272.
- [6] Chothia,C. and Lesk,A.M. (1987) J. Mol. Biol., 196, 901–917.
- [7] Grefenstette,J.J. (ed.), Proceedings of the First International Conference on Genetic Algorithms. Mahwah, NJ: Lawrence Erlbaum Associates, Inc., pp. 101–111.
- [8] Slyadnikov E. E. (2007). Physical model and associative memory of the cytoskeletal microtubule dipole system. Journal of Technical Physics, 77 (7), 77–86.

Моделювання роботи з базами даних торговельних компаній на хмарних платформах

Сергій Мінухін^{a,b}, Мирослав Башкіров^a

^a Харківський національний університет радіоелектроніки, просп. Науки. 14, Харків, 61166, Україна

^b Харківський національний економічний університет ім. С.Кузнеця, просп. Науки. 9а, Харків, 61166, Україна

Анотація

Проаналізовано проблему обробки зростаючих обсягів даних. Проведено моделювання роботи з прототипом бази даних торговельної компанії на основі згенерованих тестових даних, реалізованих на локальному ресурсі, та шляхом розгортання на сервісі Microsoft Azure. Наведено модель бази даних та опис запитів різної складності, проведений порівняльний аналіз оцінок продуктивності роботи з базами даних на локальному ресурсі та хмарному сервісі.

Ключові слова¹

Хмарні сервіси, торговельна компанія, Big Data, SQL, база даних, тестові дані, запит

1. Вступ

Збільшення об'ємів даних вимагає застосування сучасних технологій для їх зберігання та оброблення. Хмарні технології відіграють важливу роль у вирішенні завдань щодо обробки великих даних [1]. Їх використання на торговельних платформах дозволяє підвищити ефективність роботи корпоративної інформаційної системи та вдосконалити бізнес-процеси.

В дослідженні для опису предметної області та побудови моделей даних використано такі дані: товари; замовлення; філії; робітники; склади; клієнти. В процесі розроблення запитів до бази даних треба враховувати різноманіття вихідних даних для подальшої аналітики: узагальнені дані про асортимент, популярність, динаміку цін, ефективність продажів; сегментацію споживачів; ключові показники ефективності. Для заповнення таблиць великих об'ємів ефективним є генерація тестових даних - на локальному ресурсі та на сервісі Azure - з подальшим тестуванням та оцінюванням продуктивності виконання запитів.

2. Моделювання та аналіз результатів обробки даних

2.1. Створення та тестування бази даних на локальному ресурсі

На першому етапі розроблена фізична модель бази даних, яка представлена на рисунку 1, та опис таблиць бази даних з відповідною кількістю записів, що представлений у таблиці 1.

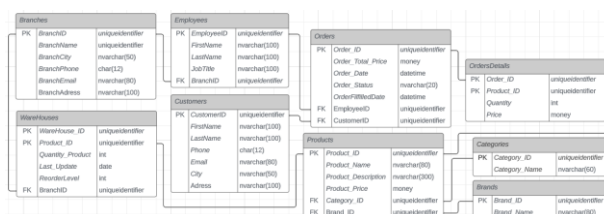


Рисунок 1: Фізична модель бази даних



Таблиця 1

Опис таблиць бази даних

Назва таблиці	Кількість записів	Об'єм (Mb)
Brands	200000	15
Categories	36	0.136
Products	2000000	273
Customers	1000000	161
Branches	100	0.136
Warehouses	1000000	76
Employees	250000	38
Orders	3000000	453
OrdersDetails	4500000	415

Для тестування створеної бази даних на локальному ресурсі розроблено SQL-запит, що відображають типові завдання для аналізу даних щодо результатів роботи торгівельної компанії: аналіз товарів за брендами і категоріями; визначення топ-продуктів за обсягами продажів; аналіз замовлень і поведінки клієнтів; аналіз показників роботи філій та інші, отримано оцінки їх реалізації.

2.2. Створення та тестування бази даних на сервісі Azure

Для реалізації другого етапу дослідження було обрано сервіс SQL Database хмарної платформи Azure [2, 3] для розгортання та оптимізації бази даних. Розроблено генератор для тестової бази даних, з використанням якого створено тестові дані таблиць вже створеної БД на локальному ресурсі в режимі on line, не використовуючи при цьому інструменти міграції бази даних з локального ресурсу. Створено запити різної складності, отримано значення метрик продуктивності щодо часу виконання запитів та завантаженості ресурсів віртуальних машин [4] з врахуванням часу генерації даних.

3. Висновки

Застосування моделювання продуктивності роботи з базами даних великих об'ємів дозволить підвищити рівень обґрунтованості прийняття рішень менеджменту компаній та рекомендацій в рамках використання сучасних аналітичних інструментів завдяки всебічному тестуванню баз даних та оцінки їх продуктивності у різних режимах розгортання.

4. Література

- [1] S.Minukhin, Performance Study Of The DTU Model For Relational Databases On The Azure Platform, Innovative Technologies and Scientific Solutions for Industries, 2022, No 1 (19), pp. 27-39. DOI: <https://doi.org/10.30837/ITSSI.2022.19.027>
- [2] Microsoft Learn. Azure SQL Database performance guidance [Електронний ресурс] – Режим доступу до ресурсу: <https://learn.microsoft.com/en-us/azure/azure-sql/database/performance-guidance?view=azuresql>
- [3] SQL Shack. Create an Azure SQL Database with built-in sample data [Електронний ресурс] – Режим доступу до ресурсу: <https://www.sqlshack.com/create-an-azure-sql-database-with-built-in-sample-data/>
- [4] SQL Performance. Azure SQL Database Performance Tuning Options [Електронний ресурс] – Режим доступу до ресурсу: <https://sqlperformance.com/2019/03/azure/azure-sql-database-performance-tuning-options>

Аналіз великих даних для виявлення туристичних трендів та персоналізації пропозицій

Даніїл Мовчан^a, Юрій Міщераков^a

^a Харківський національний університет радіоелектроніки, пр. Науки, 14, м. Харків, 61166, Україна

Abstract

Ця стаття розглядає використання великих даних для ідентифікації туристичних трендів і персоналізації пропозицій. Вона акцентує на методах Data Mining та машинного навчання для аналізу поведінки та переваг користувачів. Стаття підкреслює роль семантичного аналізу в розумінні емоційних відгуків користувачів. Основний акцент зроблено на впливі цих методів на ефективність рекомендаційних систем, спрямованих на забезпечення персоналізованого досвіду у сфері туризму.

Keywords ¹

Data Mining, аналіз великих даних, інновації в туризмі, машинне навчання, персоналізація пропозицій, поведінкові шаблони, рекомендаційні системи, семантичний аналіз, технології в туризмі, туристичні тренди.

1. Вступ

Сучасний світ туризму та подорожей характеризується швидкими змінами та великою кількістю даних, що генеруються користувачами, туристичними агентствами, соціальними мережами та іншими джерелами. Аналіз великих даних відіграє ключову роль у виявленні трендів, поведінки споживачів та створенні персоналізованих пропозицій для користувачів. Ця стаття розглядає важливість та методи аналізу великих даних у контексті туристичної індустрії.

2. Великі дані у туризмі

Впровадження великих даних у туристичній індустрії відкриває нові перспективи для аналізу та розуміння потреб та поведінки туристів. У цьому контексті великі дані можуть бути розглянуті як зібрання величезної кількості інформації з різноманітних джерел, які надають глибокий вигляд у поведінкові тенденції, уподобання туристів. Ці дані відіграють важливу роль у допомозі туристичним компаніям розробляти більш ефективні маркетингові стратегії, оптимізувати послуги та створювати персоналізовані пропозиції, які відповідають конкретним потребам і інтересам користувачів. Аналіз великих даних дозволяє фахівцям у галузі туризму виявляти закономірності та прогнозувати тренди, що є ключовими для розробки ефективних стратегій залучення та утримання клієнтів.

2.1. Джерела даних

Сектор туризму сьогодні насичений різноманітними джерелами великих даних, які включають в себе соціальні медіа, відгуки клієнтів на платформах бронювання, історію бронювань, а також дані GPS, що відстежують переміщення туристів. Соціальні мережі, такі як Instagram та Facebook, є справжнім скарбничкам інформації про тенденції та переваги

Proceedings of the *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine

EMAIL: daniil.movchan@nure.ua (A. 1); iurii.mishcheriakov@nure.ua (A. 2)

ORCID: 0009-0008-2376-1810 (A. 1); 0000-0002-5334-1808 (A. 2)



користувачів. З іншого боку, платформи відгуків, наприклад TripAdvisor, надають глибокий аналіз задоволення клієнтів та якості послуг. Ці джерела даних, разом з мобільними додатками та внутрішніми даними туристичних агентств, створюють багатий набір інформації, який може бути використаний для глибшого розуміння потреб та переваг туристів.

2.2. Обсяг даних

Зібрана інформація утворює величезний обсяг даних, що постійно зростає та оновлюється. Щоденно мільйони користувачів генерують нові дані, ділячись своїми враженнями, роблячи бронювання, пишучи відгуки, і використовуючи різні туристичні сервіси. Цей неперервний потік інформації створює виклики, пов'язані з її збором, зберіганням та обробкою. Ефективне управління цими даними вимагає застосування передових аналітичних інструментів та технологій, таких як технології обробки великих даних, що дозволяють збирати та аналізувати інформацію в реальному часі. В результаті, туристичні компанії отримують можливість швидко реагувати на зміни у поведінці та вподобаннях споживачів, забезпечуючи адаптацію своїх послуг до поточних трендів та потреб ринку.

3. Методи аналізу великих даних

Ефективний аналіз великих даних відіграє ключову роль у виявленні та розумінні складних шаблонів поведінки та переваг у сфері туризму. Інтегруючи передові технології аналізу, можна точно визначати потреби та бажання туристів, оптимізувати пропозиції та забезпечувати персоналізовані подорожі. У цьому контексті, основні методи аналізу, такі як Data Mining, машинне навчання та семантичний аналіз, відкривають нові можливості для розробки високоефективних та інноваційних рішень у туризмі, забезпечуючи унікальний досвід для кожного користувача.

3.1. Data Mining

Data Mining є ключовою технікою у розробці рекомендаційних систем, дозволяючи ідентифікувати шаблони та зв'язки у поведінці та вподобаннях користувачів [1]. Цей метод аналізу дозволяє виявляти приховані взаємозв'язки, наприклад, між популярними туристичними напрямками та особливостями, які вони пропонують. Використовуючи інформацію про попередні вибори та інтереси користувачів, рекомендаційна система може пропонувати не тільки популярні, але й індивідуально відповідні місця для відвідування, значно підвищуючи задоволеність користувачів та ефективність рекомендацій.

3.2. Машинне навчання

Методи машинного навчання надають рекомендаційним системам можливість "вчитися" з існуючих даних та покращувати свою здатність до прогнозування з часом [2]. Використовуючи алгоритми класифікації, регресії, або нейронні мережі, система може аналізувати великі обсяги інформації, включаючи поведінкові дані, пошукові запити, та відгуки користувачів, для визначення можливих вподобань та інтересів. Наприклад, система може виявити, що користувачі, які відвідують гірськолижні курорти, часто шукають також затишні готелі з каміном, та автоматично включати такі пропозиції у свої рекомендації. Це дозволяє створювати дуже точні та персоналізовані пропозиції, які постійно вдосконалюються на основі нових даних.

3.3. Семантичний аналіз

Семантичний аналіз в рекомендаційних системах використовує технології обробки природної мови для інтерпретації та розуміння людської мови у текстових даних. Це включає в

себе аналіз відгуків та коментарів користувачів для виявлення їхніх емоційних відгуків, переваг, та несподіваних вражень [3]. Наприклад, система може аналізувати відгуки на тему "романтичні вікенди" та виявити, що високо цінуються особливості, такі як приватність, унікальність місцезнаходження та якість обслуговування. Такий аналіз дозволяє не тільки підвищити точність рекомендацій, але й створити глибше розуміння нюансів індивідуальних уподобань користувачів.

4. Використання аналізу великих даних для виявлення трендів

Аналіз великих даних відіграє ключову роль у виявленні та розумінні поточних та майбутніх трендів у сфері туризму. Використовуючи різноманітні джерела даних, рекомендаційні системи можуть ідентифікувати зміни в інтересах та поведінці користувачів, що є надзвичайно цінним для адаптації туристичних пропозицій.

4.1. Виявлення популярних напрямків

В рамках рекомендаційної системи, аналіз великих даних дозволяє виявити, які з наявних у системі місць стають популярними серед користувачів. Це може включати аналіз даних про частоту пошуку певних місць, рейтинги та відгуки [4]. Наприклад, якщо система помічає зростання інтересу до певних історичних пам'яток або природних резерватів, це може вказувати на виникаючі тренди серед мандрівників. Виявлення цих трендів дозволяє системі акцентувати увагу на цих популярних напрямках, включаючи їх у перелік рекомендованих місць частіше або вище в пошукових результатах.

4.2. Адаптація до змін у перевагах споживачів

Використання великих даних також включає моніторинг та адаптацію до змін у перевагах користувачів. Це може бути зроблено шляхом аналізу даних поведінки користувачів, таких як час, проведений на перегляді певних сторінок, взаємодії з рекомендаціями, та відгуки на пропоновані місця. Наприклад, якщо дані свідчать про зростання інтересу до еко-туризму, система може реагувати на це, збільшуючи пропозиції місць, які відповідають цим критеріям, та налаштовуючи алгоритми рекомендацій, щоб вони віддавали перевагу цьому типу місць.

4.3. Прогнозування майбутніх трендів

Для підтримки довгострокової актуальності рекомендацій, аналіз великих даних може бути використаний для прогнозування майбутніх трендів. Це включає в себе вивчення історичних даних, поведінкових моделей та зовнішніх факторів, які можуть впливати на інтереси користувачів. Наприклад, з використанням аналітики та моделей прогнозування, система може визначити, що певні види подорожей або місця можуть стати більш популярними у майбутньому, враховуючи сезонність, глобальні події, або культурні тенденції. Це дозволяє системі попередньо адаптувати свої рекомендації, щоб вони відображали майбутні інтереси та переваги користувачів.

5. Персоналізація пропозицій

Персоналізація пропозицій є одним з найважливіших аспектів сучасної рекомендаційної системи. Цей підхід забезпечує, що кожен користувач отримує унікальні рекомендації, які відповідають їхнім індивідуальним перевагам та інтересам.

5.1. Індивідуальний підхід до користувача

Персоналізація починається з детального розуміння уподобань та інтересів кожного користувача. Це може бути досягнуто шляхом аналізу попередніх пошуків, бронювань та відгуків користувача. Наприклад, якщо користувач часто шукає морські курорти або віддає перевагу готелям вищого класу, система автоматично адаптує свої рекомендації, щоб відобразити ці переваги. Це забезпечує, що кожна пропозиція має вищу ймовірність відповідати очікуванням користувача.

5.2. Покращення задоволення клієнтів

Персоналізація також спрямована на підвищення загального задоволення користувачів. Це досягається не тільки шляхом відповідності їхніх очікувань, але й шляхом запропонування унікальних, неочікуваних варіантів, які можуть зацікавити користувачів. Наприклад, система може запропонувати маловідомі, але високо оцінені туристичні місця, які відповідають загальним інтересам користувача, здивувавши його новизною та оригінальністю вибору.

5.3. Динамічна адаптація пропозицій

Персоналізація пропозицій у рекомендаційних системах також вимагає динамічної адаптації до змін у вподобаннях користувачів. Це означає, що система постійно оновлює свої рекомендації на основі останньої поведінки користувача. Наприклад, якщо користувач починає шукати інформацію про культурні події або музеї, система швидко включає такі місця у свої пропозиції.

6. Висновки

Аналіз великих даних в індустрії туризму не тільки сприяє виявленню трендів та адаптації до змін у вподобаннях користувачів, але й дозволяє створювати високо персоналізовані та ефективні пропозиції. Це підвищує задоволеність клієнтів та сприяє зростанню туристичного бізнесу. Застосування інноваційних технологій та методів аналізу даних є ключовими для досягнення успіху в цій динамічній галузі.

7. Перелік джерел посилання

- [1] Han, J., Kamber, M., & Pei, J. Data Mining: Concepts and Techniques, Morgan Kaufmann, 2011.
- [2] Alpaydin, E. Introduction to Machine Learning, MIT Press, 2020.
- [3] Manning, C. D., & Schütze, H. Foundations of Statistical Natural Language Processing, 1999.
- [4] Liu, B. Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data. Springer, 2011.

Дослідження Методів Машинного Навчання для Прогнозування

Аліна Овчаренко^a та Роксана Петрова^a

^a Харківський Національний Університет Радіоелектроніки, 14, просп. Науки., Харків, 61166, Україна

Анотація

Дана робота розглядає системний підхід до аналізу та моделювання даних, а також використання методів машинного навчання для побудови точних прогнозів. Дослідження проводились на прикладі прогнозуванні врожайності, але можуть бути застосована для будь-якої галузі. Проведені дослідження та практична реалізація системи прогнозування мають значущий потенціал для підвищення продуктивності та стійкості сільськогосподарського виробництва.

Ключові слова ¹

Аналіз, врожайність, випадкові ліса, дерева рішень, дослідження, машинне навчання, прогнозування

1. Вступ

Наша країна має довгу традицію сільськогосподарського виробництва та є однією зі значущих аграрних держав у світі. Забезпечення стабільної врожайності та ефективного сільського господарства є ключовим завданням для забезпечення продовольчої безпеки та економічного розвитку України.

Наявність точних та актуальних прогнозів врожаю [1] є критично важливим для планування та оптимізації процесів сільського господарства. Дослідження та впровадження інноваційних методів для прогнозування врожаю та оптимізації виробництва стають все важливішими завданнями в цій галузі.

2. Хід дослідження

Робота присвячена впровадженню методів машинного навчання [2] для прогнозування врожайності сільськогосподарських культур. Основною метою є створення інформаційної системи [3], яка дозволить аграрним підприємствам та фермерам отримувати точні та актуальні прогнози врожаю. Використання інноваційних підходів дозволить підвищити якість прогнозів та зменшити ризики в сільському господарстві.

В ході виконання проведено дослідження предметної області, охоплюючи аналіз проблем та визначення сфери застосування методів машинного навчання для прогнозування врожайності сільськогосподарських культур. Сформовано постановку задачі, орієнтовану на використання машинного навчання для покращення прогнозування урожайності та оптимізації агропромислового сектора.

Для розробки моделей прогнозування чітко визначення та докладний аналіз кожного параметра вхідних даних – фундамент, необхідний для досягнення високої точності та

Proceedings of the *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine

EMAIL: alina.ovcharenko1@nure.ua (A. 1); roksana.petrova@nure.ua (A. 2)

ORCID: 0000-0001-5886-8943 (A. 2)



ефективності в прогнозуванні врожайності сільськогосподарських культур. Використання офіційних [4] даних з врожайності минулих років при машинному навчанні є необхідним етапом для тренування точних та надійних моделей прогнозування урожайності. Окрім цього, у роботі при прогнозуванні було вирішено звернути увагу на параметри [5], на які людина зазвичай не може вплинути. Ці зовнішні фактори включають в себе кліматичні умови, географічне розташування та погодні умови, тип ґрунту. Для збору цих даних було вирішено звернутися до карти ґрунтів України [6] та архівів щоденників погоди [7].

Розглядаючи ці різноманітні підходи, ми можемо краще розуміти їх сильні та слабкі сторони, а також визначити, чому вибір певного методу, наприклад, методу випадкових лісів, є обґрунтованим для конкретного випадку прогнозування врожайності. Тому перед виконанням практичної частини задачі прогнозування врожайності було вирішено дослідити різні підходи у сфері машинного навчання, зокрема в контексті прогнозування урожайності.

Обрано методи машинного, такі як метод часових рядів [8], лінійна регресія [9], метод опорних векторів [10] та випадкових лісів [11]. Проведено порівняльний аналіз цих методів, визначено їх переваги та недоліки. Враховуючи необхідність роботи з великим обсягом даних в котрих можливі непередбачувані випадки (викиди) [12] та все ще необхідний певний рівень точності, було зосереджено увагу на методі випадкових лісів. Однак, враховуючи недоліки метода було вирішено розглянути можливості вдосконалення його застосування, включаючи додавання ваг до прикладів.

3. Висновки

Використання методів машинного навчання є багатообіцяючим напрямком для покращення прогнозування врожайності у сільському господарстві. Вибір конкретного методу повинен бути обґрунтованим з урахуванням особливостей задачі та доступних даних. Розгляд вагомості кожного методу та його можливостей у контексті покращення сільськогосподарської продуктивності є ключовим аспектом подальших досліджень та впровадження в практику.

4. Перелік джерел посилання

- [1] О. І. Зінченко, Програмування врожайності сільськогосподарських культур, Редакційно-видавничий відділ Уманського національного університету садівництва, Умань, 2015.
- [2] N. J. Nilsson , Introduction to machine learning, Stanford, CA 94305, 2005.
- [3] В. Ф. Ситник, Т. А. Писаревська, Н. В. Єрьоміна, Основи інформаційних систем, 2-е вид., Київський національний економічний університет, Київ, 2001.
- [4] Державна служба статистики України. URL: <https://www.ukrstat.gov.ua>.
- [5] Л. Ю. Забродоцька, Основи агрономії, Інформаційно-видавничий відділ Луцького національного технічного університету, Луцьк, 2019.
- [6] Карта ґрунтів України. URL: <https://superagronom.com/karty/karta-gruntiv-ukrainy>.
- [7] Архів погоди. URL: https://www.meteoblue.com/uk/weather/historyclimateweatherarchive/kher son_ukraine_706448?fcstlength=1m&year=2022&month=11.
- [8] G. Shmueli, K. C. Lichtendahl Jr, Practical time series forecasting with R a hands-on guide, 2nd ed. axelrod schnall publishers, 2016.
- [9] T. Hastie, R. Tibshirani , J. Friedman, The elements of statistical learning data mining, inference, and prediction, 12th ed., Springer, 2017.
- [10] A. Ben-Hur, Support vector clustering, Vol. 2 of Journal of machine learning research, 2001. pp. 125–137.
- [11] L. Breiman, Random forests. Statistics Department University of California, Berkeley, CA 94720, 2001.
- [12] С. М. Іванов, Н. К. Максишко, Д. О. Бречко, Інтелектуальний аналіз даних, Запорізький національний університет, Запоріжжя, 2020.

Аналіз якості обслуговування в сенсорних мережах

Ігор Пастушенко^а, Денис Черненко^а

^аХарківський національний університет радіоелектроніки, пр. Науки. 14, Харків, 611166, Україна

Анотація

Проведено визначення основних характеристик якості обслуговування безпроводових сенсорних мереж (БСМ). Показано взаємозв'язок характеристик якості обслуговування з різними факторами реалізації мережевої моделі. Показано залежність якості обслуговування сенсорних мереж від особливостей протоколів маршрутизації. Проведено порівняльний аналіз протоколів ТМАС, ZigBeeMAC з GTS та BaselineMAC. Наведено відповідні критерії якості обслуговування.

Ключові слова¹

Безпроводова сенсорна мережа, характеристики якості обслуговування, протоколи маршрутизації

1. Вступ

Безпроводові сенсорні мережі (БСМ) з кожним роком знаходять ширше застосування в різних галузях промисловості та сферах діяльності людини. Вони широко використовуються не тільки для збору й обробки даних та керування промисловими об'єктами, але й у звичайному побуті. Це робить її одним з найперспективніших видів технологій, які вже сьогодні називають «технології майбутнього» [1].

Якість обслуговування (Quality of service - QoS) вводиться в мережах для визначення здатності мережі відповідати заданим вимогам і характеристикам. Якість обслуговування визначає, чи може мережа надати необхідний сервіс передачі даних за заданих умов.

У класичних мережах як такі параметри використовують смугу пропускання, затримки при передачі пакетів, розкид затримки, а також можливість доставки пакета [2].

Якість обслуговування в безпроводових сенсорних мережах – це інтегральна характеристика, яка визначає можливість мережі виконувати збір, обробку та передачу даних за тих чи інших заданих умов. Характерні для безпроводових сенсорних мереж особливості, що визначаються умовами їх функціонування, що впливають на визначення якості обслуговування, наступні: обмежені ресурси, стійкість, різні види трафіку, модель передачі, мобільність [3].

На відміну від класичних мереж безпроводові сенсорні мережі надають сервіс не тільки передачі даних, але також і по збору та їх обробці. Відповідно якість обслуговування в безпроводових сенсорних мережах і його показники також відрізнятимуться від класичного уявлення і потребують аналізу.

2. Аналіз якості обслуговування у сенсорних мережах

До характеристик якості обслуговування безпроводових сенсорних мережах відносять такі параметри як: затримка, пропускна здатність, втрати, час життя мережі, покриття заданої області, стійкість до зміни топології [4].

Параметри якості обслуговування часто пов'язані між собою. У табл.1 наведені фактори, що впливають на характеристики якості обслуговування, що відносяться до того чи іншого рівня мережевої моделі.

Proceedings of the *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine
EMAIL: ihor.pastushenko@nure.ua (A. 1); denys.chernenko@nure.ua (A. 2)



Таблиця 1

Взаємозв'язок характеристик якості обслуговування з різними факторами реалізації мережевої моделі.

Характеристика якості обслуговування	Фізичний рівень	Канальний рівень	Мережевий рівень
Затримка	Спосіб кодування	Розклад доступу до каналу	Час визначення маршруту, довжина колії
Пропускна здатність	Розмір повідомлень	Синхронізація доступу до каналу, надлишкові пакети	Маршрутизація по кількох шляхах, додаткові дані маршрутизації в пакеті, кластеризація
Втрати	Рівень шуму	Колізії	Тупикові маршрути, зациклювання
Час життя мережі	Спосіб кодування, потужність передачі	Час активного/пасивного режимів, повторна передача пакетів	Використання тих самих вузлів при побудові маршруту, кластеризація
Покриття	Потужність передачі	-	-
Стойкість	Потужність передачі	Період зміни режимів	Динамічна маршрутизація, кластеризація

Забезпечення якості обслуговування залежить від особливостей протоколу на кожному рівні мережевої моделі: фізично якість обслуговування залежить від рівня шуму, способу кодування та ін., на канальному – від часу доступу до середовища та надійності доставки пакетів. Залежно від реалізації алгоритму маршрутизації можуть змінюватися витрати енергії мережі, затримка та швидкість доставки повідомлень.

Для дослідження параметрів якості обслуговування сенсорних безпроводових мереж використовувався симулятор безпроводових сенсорних мереж Castalia 3.2. При моделюванні задані такі параметри: час симуляції – $t_{sim} = 51c$, з якої $1c$ відводиться для налаштування вузлів, розмір пакету $sp = 100$ байт, розмір заголовка пакета $sh = 5$ байт, пропускна здатність каналу зв'язку – 1024 КБ/с. Під час дослідження проводилося порівняння трьох MAC-протоколів – T-MAC, ZigBeeMAC та BaselineMAC. Результати моделювання представлено на рис.1, 2.

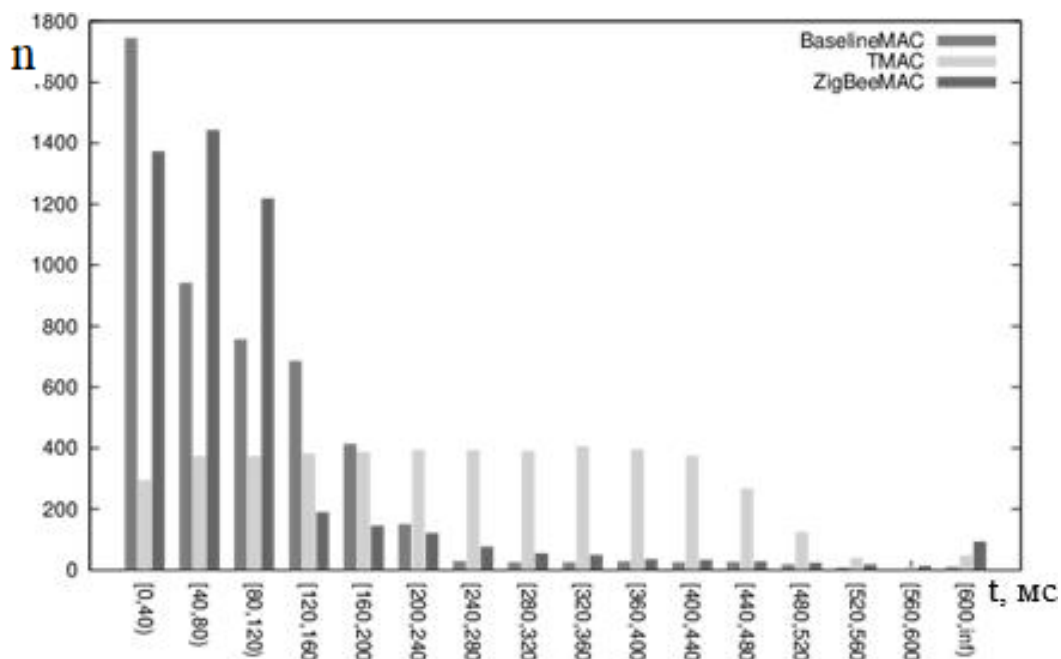


Рисунок 1: Гістограма кількості доставлених пакетів за різні проміжки часу для різних MAC-протоколів

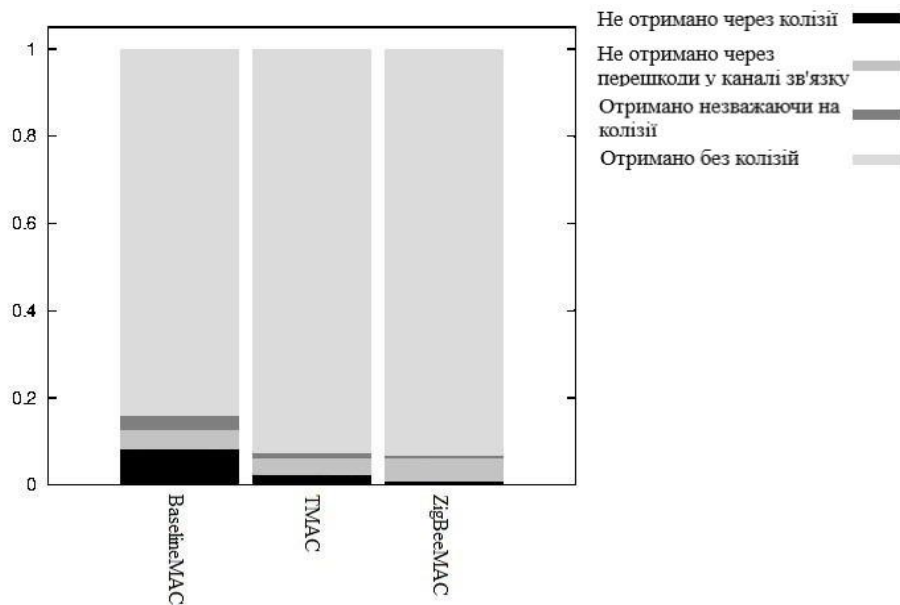


Рисунок 2: Відношення кількості доставлених пакетів до втрачених різних MAC-протоколів

3. Висновки

Проведено визначення основних характеристик якості обслуговування БСМ: затримка, пропускна здатність, втрати, час життя мережі, покриття заданої області, стійкість до зміни топології. Показано взаємозв'язок характеристик якості обслуговування з різними факторами реалізації мережевої моделі. Наведено критерії, що стосуються передачі даних у БСМ, за якими можна оцінювати якість обслуговування.

Показано залежність якості обслуговування сенсорних мереж від особливостей протоколів маршрутизації. Порівнювалися протоколи TMAC, ZigBeeMAC з GTS та BaselineMAC. BaselineMAC і ZigBeeMAC конкурують між собою за затримкою доставки, більшість пакетів ZigBee доставляє в межах 120 мс. У TMAC затримка розподілена практично рівномірно в межах півсекунди. Проте якщо важливою є швидкість доставки, застосовувати цей протокол не є доцільним. BaselineMAC – швидкий, але не кращий за надійністю протокол.

4. References

- [1] Л.І. Мельнікова, О.В. Лінник, М.В. Кривошапка, В.О. Барсук, Застосування евристичної процедури багатокритеріальної оптимізації до вибору варіанта мовного кодеку в IP-мережі. Проблеми телекомунікацій 26(2020) 23-32. doi: 10.30837/pt.2020.1.02
- [2] L. Melnikova, E. Linnyk, D. Ageyev, O. Melnikova, N. Kryvoshapka and V. Barsouk, Minimizing the Route of Sink Node in Wireless Sensor Network, 2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology, PIC S&T, Kyiv, Ukraine, 2019, pp. 861-864, doi: 10.1109/PICST47496.2019.9061563
- [3] X. Dong, M. C. Vuran, and S. Irmak, Autonomous precision agriculture through integration of wireless underground sensor networks with center pivot irrigation systems. Ad Hoc Networks 11(7) 1975–1987. doi: 10.1016/j.adhoc.2012.06.012.
- [4] D. Hemanand, C. Senthilkumar, Omar S. Saleh, B. Muthuraj, A. Anand, V. Velmurugan, Analysis of power optimization and enhanced routing protocols for wireless sensor networks. Measurement: Sensors 25(2023) doi: 10.1016/j.measen.2022.100610.
- [5] Yuanli Wang, Xianghui Liu and Jianping Yin, Requirements of Quality of Service in Wireless Sensor Network, International Conference on Networking, International Conference on Systems and International Conference on Mobile Communications and Learning Technologies ICNICONSMCL'06, Morne, Mauritius, 2006, pp. 116-116, doi: 10.1109/ICNICONSMCL.2006.185.

Аналіз засобів автоматизації визначення діагнозу для інформаційної системи "Медичний центр"

Геннадій Перегуда^a, Антон Міщераков^a

^a Харківський національний університет радіоелектроніки, пр. Науки, 14, м. Харків, 61166, Україна

Abstract

Робота присвячена опису розробки інформаційної системи для автоматизації визначення діагнозів у медичних центрах. Основна мета полягає в створенні інструменту, який дозволить пацієнтам оцінювати ймовірність наявності хвороби, вводячи власні симптоми та індивідуальні характеристики. Описаний проект спрямований на покращення точності визначення діагнозів та полегшення роботи медичного персоналу, щоб забезпечити якісне медичне обслуговування та підвищити свідомість пацієнтів.

Keywords¹

Медичний центр, інформаційна система, діагноз, визначення діагнозу, автоматизація, хвороба, симптом, індивідуалізована медицина, інформаційні технології у здоров'ї, автоматизація в медицині, поліпшення здоров'я пацієнтів, пацієнтська самодіагностика

1. Вступ

У наш час інформаційні технології надають суттєвий вплив на різні сфери життя, включаючи медичну сферу. Зростання обсягу медичної інформації та швидкість її обробки створюють нові можливості для вдосконалення процесів діагностики та лікування хвороб. У цьому контексті робота, присвячена розробці інформаційної системи для визначення ймовірності наявності певної хвороби на основі введених симптомів та індивідуальних характеристик пацієнта, набуває особливої актуальності.

В умовах постійного розвитку медичної науки та зростання обсягу медичних даних лікарі та пацієнти стикаються із збільшеною складністю прийняття обґрунтованих рішень. Інформаційна система, яку пропонується розробити, вирішує конкретну проблему: надання пацієнтам можливості самостійно оцінювати свій стан здоров'я та отримувати інформацію про можливі діагнози на основі введених даних.

Важливість цього проекту визначається потребою в покращенні доступності медичної інформації для громадськості та надання пацієнтам засобів для більш свідомого управління власним здоров'ям. Така інформаційна система не тільки спростить процес самодіагностики для користувачів, а також може стати інструментом для та раннього виявлення захворювань.

Отже, ця робота має реальну практичну значущість, оскільки вона спрямована на створення інноваційної інформаційної системи, яка може вплинути на підвищення рівня інформованості та відповідальності пацієнтів у відношенні їхнього здоров'я.

2. Постановка завдання

Центральною проблемою, яку вирішує ця робота, є розробка інформаційної системи, спрямованої на підвищення доступності та точності визначення діагнозів для пацієнтів. Основне

Proceedings of the *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine

EMAIL: hennadii.perehuda@nure.ua (A. 1); anton.mishcheriakov@nure.ua (A. 2)

ORCID: 0009-0006-5043-4917 (A. 2)



завдання полягає в тому, щоб створити засіб, який дозволить користувачам, засновано на їхній власній інформації, оцінювати ймовірність наявності певної хвороби.

Підходи до вирішення проблеми:

- розробка комплексних медичних інформаційних систем (МІС). Об'єднання аспектів обробки медичних даних, від збору до візуалізації, для автоматизації та спрощення роботи медичного персоналу;
- використання систем штучного інтелекту (ШІ). Впровадження методів машинного навчання та глибоких нейронних мереж для автоматизованого визначення діагнозу, враховуючи різні фактори;
- стандартизація медичних даних та клінічних рекомендацій. Розвиток та використання стандартів для однорідності та узгодженості визначення діагнозів.

Таким чином, розробка інформаційної системи, що враховує індивідуальні характеристики пацієнта для точного визначення ймовірності хвороби, відповідає насущним потребам сучасної медицини та пацієнтів. Результатом роботи має бути ефективний інструмент, спрямований на поліпшення якості медичного обслуговування та підвищення обізнаності пацієнтів щодо їхнього здоров'я.

3. Підходи до вирішення задачі

В даній роботі проблема визначена як необхідність розробки інформаційної системи, яка дозволяє пацієнтам самостійно оцінювати ймовірність наявності хвороби на основі введених симптомів. Головним завданням є створення інформаційної системи, яка дозволить пацієнтам оцінювати ймовірність наявності хвороби на основі введених симптомів. Результатом роботи має бути інструмент, який може покращити здоров'я та самопізнання пацієнтів, а також сприяти підвищенню обізнаності у галузі медицини серед громадськості.

Для вирішення цієї проблеми, розробляється інформаційна система, яка базується на обширних медичних даних та експертних знаннях для визначення ймовірності наявності хвороби на основі симптомів, вказаних пацієнтом. Система має враховувати індивідуальні характеристики пацієнта, такі як вік, стать, хронічні хвороби, погані звички, що має дозволити робити більш індивідуалізовані та точні діагнози. Такий підхід дозволить вирішити недоліки існуючих систем, забезпечуючи більш ефективний та зручний інструмент для пацієнтів у визначенні свого стану здоров'я.

Нижче зазначено основні етапи процесу визначення діагнозу:

- збір інформації від пацієнта. Система дозволяє пацієнтам вводити інформацію про симптоми, вік, стать, хронічні хвороби, погані звички та попередні травми, що є ключовим етапом для аналізу та визначення діагнозу;
- аналіз симптомів та індивідуальних характеристик. Система впроваджує алгоритм для обробки введених даних та аналізу симптомів, враховуючи індивідуальні характеристики пацієнта, що є важливим для точного встановлення діагнозу;
- визначення ймовірності хвороби. Система розраховує ймовірність наявності певної хвороби на основі аналізу симптомів та інших індивідуальних факторів, виражаючи її у відсотках для зручності користувача;
- візуалізація результатів. Результати аналізу представляються системою зрозуміло та доступно, включаючи інформацію про ймовірність наявності різних хвороб та рекомендації для пацієнта;
- безпека і конфіденційність даних. Система має забезпечувати високу безпеку особистої інформації пацієнтів, зокрема шляхом застосування відповідних заходів шифрування та конфіденційності обробки даних;
- можливість розширення системи. Система розроблюється з урахуванням можливості легкого розширення функціональності та додавання нових хвороб і симптомів для аналізу, забезпечуючи гнучкість та актуальність;
- загальна мета та потенційні переваги системи. Загальна мета полягає в створенні інструменту для інформованого самостійного оцінювання стану здоров'я пацієнтів. Система пропонує оцінку ймовірності наявності хвороби у процентах, враховуючи більше

індивідуальних характеристик, що дозволить отримати більш точні та персоналізовані результати.

Отже, розробка цієї інформаційної системи вирішує ключові проблеми, описані в постановці завдання, надаючи інноваційний інструмент для пацієнтів та сприяючи покращенню їхнього самопізнання та обізнаності у галузі медицини.

4. Висновки

У цій роботі визначено та проаналізовано проблеми, пов'язані з встановленням діагнозу в медичних інформаційних системах. Головним завданням було формулювання вимог до розробки інформаційної системи, яка дозволить пацієнтам самостійно оцінювати ймовірність наявності хвороби на основі наданих симптомів.

Отриманий інструмент виступає не лише як засіб для поліпшення здоров'я та самопізнання пацієнтів, але й як ініціатива, спрямована на підвищення рівня обізнаності у сфері медицини серед широкої маси населення.

Ця робота не лише розширює розуміння процесів визначення діагнозу в медичних інформаційних системах, але й вносить практичний внесок у сферу ефективного взаємодії пацієнтів із сучасною медициною. Реалізація цього проекту може значно полегшити доступ до медичної інформації, поліпшити якість обслуговування та допомогти у вчасному виявленні та лікуванні хвороб.

5. References

- [1] Гаген В. Розробка методів аналізу структурованих медичних записів для системи обміну медичною інформацією. *Measuring and computing devices in technological processes*. 2022. № 1. С. 49–56. URL: <https://doi.org/10.31891/2219-9365-2022-69-1-7> (дата звернення: 24.11.2023).
- [2] Годованець Б. В., Нодованець В. Створення прототипу інформаційної системи для медичних закладів : master's thesis. 2019. URL: <http://elartu.tntu.edu.ua/handle/lib/31145> (дата звернення: 24.11.2023).
- [3] Ключко О. М. Медична інформаційна система моніторингу стану здоров'я населення із захистом персональних даних. *Medical Informatics and Engineering*. 2020. № 1. С. 17–28. URL: <https://doi.org/10.11603/mie.1996-1960.2020.1.11126> (дата звернення: 24.11.2023).
- [4] Подоляко В. І. Інформаційна система підтримки діяльності медичного центру : master's thesis. 2021. URL: <https://essuir.sumdu.edu.ua/handle/123456789/86876> (дата звернення: 24.11.2023).
- [5] Рогова Є. І. Теоретичні основи правового забезпечення інформаційної безпеки. Актуальні проблеми держави і права. 2020. № 86. С. 190–196. URL: <https://doi.org/10.32837/apdp.v0i86.2436> (дата звернення: 24.11.2023).
- [6] Швець Ю. Адміністративно-правові основи захисту лікарської таємниці: досвід європейських країн та перспективи для України. *Юридичний вісник*. 2020. № 3. С. 241–247. URL: <https://doi.org/10.32837/yuv.v0i3.1947> (дата звернення: 24.11.2023).
- [7] Юдкова К. В. Особливості визначення поняття "інформаційна система". *Інформація і право*. 2015. № 2 (14). С. 39–44.

Дослідження Використання Неактуальних Практик як Проблеми Підтримки CSS-коду

Наталя Понікаровська^a, Марія Широкопетлева^a

^a *Kharkiv National University of Radio Electronics, 14 Nauky Avenue, Kharkiv, 61166, Ukraine*

Abstract

Дослідження виявляє проблеми у підтримці CSS-коду, зумовлені застосуванням неактуальних практик його написання. Пропонується рішення через рефакторинг та використання відповідних методологій CSS.

Keywords ¹

Technical debt, CSS optimization, CSS development methodology

1. Вступ

У світі програмування існує парадокс: розробники усвідомлюють, що безлад у коді сповільнює їх роботу, але, відчуваючи тиск дедлайнів та кваплячись, знов створюють безлад. “Ім бракує часу, щоб працювати швидко!” [1, с. 26]. Правда в тому, що лише постійна увага до чистоти в коді може дозволити одночасно дотримуватися графіка та працювати ефективно.

2. Постановка задачі

В сучасному світі важко уявити повсякденне життя без веб-технологій. Навчання, робота, хобі, спілкування - на все знайдеться відповідний веб-ресурс. Від простих html-сторінок веб розвинувся у глобальну систему інтерактивних веб-додатків. Такі фактори як зростання складності дизайну, різноманіття пристроїв, поява нових ідей, що призводить до змін у вимогах, ведуть до постійних змін у веб-інтерфейсах, що часто призводить до появи коду з високим рівнем технічного боргу. Беручи такий проект у роботу часто важко зрозуміти з чого почати та як відладити проект повністю його не переписавши. У даній роботі пропонуємо розглянути проблеми, що виникають у процесі підтримки css-коду з високим рівнем технічного боргу

3. Опис дослідження

Основне питання дослідження - це визначення певних ознаках технічного боргу у css-коді [2], яке розглянемо на декількох прикладах використання застарілих та неактуальних практик.

Перша проблема - велика вкладена вага (рис.1а). Поява препроцесорів значно полегшила життя веб-розробникам, бо стала можлива інкапсуляція стилів. Але зловживання вкладеністю селекторів може значно ускладнити розуміння та редагування стилів та призводить до повторення (замість перевикористання) css-правил, появи надлишкових класів.

Глобальні селектори (приклад наведено на рис.1б) впливають на всі елементи в документі, що підвищує ризик виникнення перевизначень або необхідність створення зайвих класів, що будуть перекривати глобальні стилі. Звертатись до селекторів таким чином припустимо тільки у правилах для скидання стилів.

Proceedings of the *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine

EMAIL: natalia.ponikarovska.cpe@nure.ua (A.1), marija.shirokopetleva@nure.ua email1@mail.com (A.2)

ORCID: 0009-0000-5243-3287 (A. 1); 0000-0002-7472-6045 (A. 2)





Рисунок 1 - Приклади: а) надлишкової вкладеності; б) глобальних селекторів; в) недоцільного використання css-правил; г) повторення css-правил

Ознакою похибок у проектуванні є використання директиви !important, що змінює природний порядок спадкоємності стилів. Наявність даної директиви також може бути показником невідповідності UI бібліотеки, що використовується, реальним потребам продукту. Також використання інлайн-стилів ускладнюють підтримку коду, роблячи його менш читабельним та гнучким для змін.

Застосування CSS-правил, що не призначені для вирішення певної задачі може призводити до складнощів у підтримці та адаптації до сучасних вимог та технологій: (див. рис.1в).

Наступна проблема - відсутність хелперів та готових (або розроблених под потреби проекту) “сіток” (рис.1г). Повторення css-правил ускладнює обслуговуваність коду (бо у разі потреби внесення змін доведеться робити це в кількох місцях) та значно збільшує його об’єм.

Отже, якщо ми маємо справу з css-кодом, що має одну чи декілька перелічених ознак технічного боргу, то такий код потребує рефакторингу. Першим кроком доречно узгодити методологію написання CSS-коду - OOCSS, Atomic CSS тощо (або розробити власну, підпорядковану реальним потребам проекту, методологію) та перевірити css-код на якість та унікальність. Для цього треба позбутися інших ознак технічного боргу, таких як велика вкладена вага, відмовитись від використання глобальних селекторів, перевірити код на директиву “!important” та позбутись причин її використання, винести інлайн-стили у css-документи, перевірити чи застосовані правила є призначеними для вирішення певних задач, написати хелпери та винести кольори й все, що необхідно та доречно, у змінні.

4. Висновки

В даній роботі розглянуло деякі ознаки технічного боргу у css-кодї. Наводяться приклади css-коду, які можуть призводити до складнощів у збереженні консистентності та адаптації до змін. Дослідження визначає проблеми і рекомендує рефакторинг як спосіб вирішення перелічених проблем. Визначається важливість узгодження методології написання css-коду. Узагальнюючи, дослідження вказує на необхідність вдосконалення практик розробки css-коду з метою підвищення його читабельності, підтримки та адаптабельності до змін.

5. References

[1] Чистий код: створення і рефакторинг за допомогою Agile / пер. з англ. І.Бондар-Терещенко. – Харків : Вид-во “Ранок” : Фабула, 2021. – 448 с.
 [2] Verou L. CSS Secrets: Better Solutions to Everyday Web Design Problems / Lea Verou., 2015. – 354 с. – (Third Edition).

Програмне моделювання лінгвістичного представлення детермінованих графів

Володимир Слесарев^a, Олексій Сенченко^b, Микола Притула^{a,b}

^a Національний ТУ «Дніпровська політехніка», пр. Дмитра Яворницького, 19, Дніпро, 49005, Україна

^b Інститут прикладної математики і механіки НАН України, вул. Батюка, 19, 84116, Україна

Анотація

Наведено аспекти реалізації програмної моделі лінгвістичного представлення детермінованих графів. Окреслено переваги наведеної реалізації.

Ключові слова¹

Детермінований граф, представлення, моделювання

1. Вступ

Однією з важливих проблем у сучасній математиці та моделюванні є зручне задання як скінченних, так і нескінченних об'єктів. Одним із ефективних способів такого задання є представлення об'єкту за допомогою твірних елементів та співвідношень між цими елементами. У такий спосіб можна задавати групи, напівгрупи, автомати; відомий ряд теоретичних та прикладних задач, які безпосередньо пов'язані із таким способом задання. Наразі таке представлення розповсюджено [1] на детерміновані [2] графи, які можуть бути ефективно використані в якості середовища у багатьох прикладних задачах. При цьому, детерміновані графи задаються словами в алфавіті міток їх вершин, тому можна говорити саме про лінгвістичне представлення детермінованих графів. Оскільки на даний час графи є концептуальним інструментом технічної документації та, крім того, є широко вживаною програмною абстракцією, тому важливу роль відіграє програмне моделювання результатів з метою їх підтвердження, аналізу, уточнення та удосконалення практичною реалізацією. Крім того, можливе застосування лінгвістичного представлення детермінованих графів у складних інженерних системах потребує розробки програмних прототипів алгоритмів, які запропоновано у цьому представленні, з метою подальшого аналізу можливості їх впровадження в промислові стандарти.

2. Основні означення

Розглядаються неорієнтовані, скінченні, непорожні, зв'язні прості графи з розміченими вершинами $G = (V, E, X, \xi(V))$, де V – множина вершин графа, E – множина його ребер, $\xi(V): V \rightarrow X$ – всюдिवизначена функція розмітки вершин графа символами скінченного алфавіта $X = \{x_1, \dots, x_p\}$. Розмічений граф G називається детермінованим (Д-графом) [2], якщо всі вершини у відкритому околі кожної його вершини мають різні мітки: $\forall v, v_1, v_2 \in V (v_1, v_2 \in E(v) \wedge \xi(v_1) = \xi(v_2)) \rightarrow v_1 = v_2$. Шлях $p = v'_1 \dots v'_k$ розглядаємо як послідовність $\xi(p) = \xi(v'_1) \dots \xi(v'_k)$ міток вершин, що входять до p . Зауважимо, що у випадку, коли відома карта графа та початкова вершина шляху, то за послідовністю міток $\xi(p)$ однозначно відтворюється послідовність вершин, що входять до шляху, тому Д-графи є дуже зручними для їх дослідження за допомогою агентів, що переміщуються всередині графу. Зафіксуємо вершину $v_0 \in V$, яку будемо називати ініціальною, далі вважаємо, що всі шляхи в графі починаються з ініціальної

Proceedings of the *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine

EMAIL: sliesariiev.v.v@nmu.one (A. 1); senchenko.a76@gmail.com (A. 2); elanir358@gmail.com (A. 3)

ORCID: 0000-0002-5385-7118 (A. 1); 0000-0002-6684-6530 (A. 2); 0000-0001-8426-6364 (A. 3)



вершини. Шлях (слово) $p = x'_1 \dots x'_k$ назвемо припустимим для вершини $v'_1 \in V$, якщо $\xi(v'_1) = x'_1$, та існують такі вершини $v'_2 \dots v'_k \in V$, що $\xi(v'_2) = x'_2, \dots, \xi(v'_k) = x'_k$ і $(v'_1, v'_2), \dots, (v'_{k-1}, v'_k) \in E$.

У [1] було розроблено представлення Д-графа $G = (V, E, X, \xi(V), v_0)$ парою $\{C, L\}(x')$ скінченних множин слів в алфавіті X , для якої слова множини C описують цикли графа G , а слова L описують його висячі вершини. Множини C та L називають компонентами пари $\{C, L\}(x')$. Було висунуто вимоги, за виконанням яких Д-граф $G = (V, E, X, \xi(V), v_0)$ вважається представленням за парою $\{C, L\}(x')$:

- а) $\xi(v_0) = x'$;
- б) всі слова з C та L є припустимими для вершини v_0 ;
- в) кожне слово $r \in C$ визначає у G цикл $v_0 r = v_0$;
- г) для кожного слова $q \in L$ у G вершина $v_0 q$ є висячею;
- д) для кожної висячої вершини $v \in V$, що є відмінною від ініціальної вершини v_0 , існує хоча б одне таке слово $q \in L$, що $v_0 q = v$;
- е) для будь-якої пари $\{C, L\}(x')$ або не існує представлення, або це представлення визначається однозначно.

При цьому декільком різним парам може відповідати одне й те ж представлення.

У [1] було визначено алгоритм (АП), який за заданою парою $\{C, L\}(x')$ або будує Д-граф $G(\{C, L\}(x'))$, який є представленням за нею, або показує, що за цією парою не існує представлення, тобто не існує такого Д-графа, що відповідає умовам (а) – (е). Цей алгоритм у якості допоміжного використовує алгоритм редукції (АР), який спеціальним способом однозначно перетворює довільний граф G у детермінований. Пара $\{C, L\}(x')$ називається визначальною для графа G , якщо $G(\{C, L\}(x')) \cong G$. Також у [1] визначено особливу канонічну визначальну пару для будь-якого Д-графа G . Для цього на множині всіх слів в алфавіті X введено лінійний порядок \ll , у відповідності з яким побудовано найкоротше кістякове дерево графа G , за допомогою якого із використанням алгоритму АК побудована ця канонічна визначальна пара $\{\Sigma_G, \Lambda_G\}$. Було знайдено деякі метричні характеристики пари $\{\Sigma_G, \Lambda_G\}$.

3. Програмна модель

В якості технічної специфікації для програмного моделювання алгоритмів АП, АР та АК визначено роботу [1]. Технічна реалізація [3] виконана з використанням мови програмування Python та бібліотеки для моделювання мереж NetworkX [4], з дотриманням вимог специфіки статичного та синтаксичного аналізатора MyPy та Pylint. Якість коду, згідно цих перевірок, програмно оцінена як 9.3 / 10 за стандартом PEP8. Ця оцінка не є максимально можливою завдяки наявності в функціях збільшеної кількості, локальних змінних, а також використання назв змінних іменованих однією літерою, що є не рентабельним при написанні складних програмних модулів, але ця реалізація обрана свідомо, для відповідності термінам, використаним у роботі [1]. На рис. 1 наведено результати оцінки коду:

```
***** Module alglib
alglib.py:19:34: C0103: Argument name "G" doesn't conform to snake_case naming style (invalid-name)
alglib.py:114:4: C0103: Variable name "G_" doesn't conform to snake_case naming style (invalid-name)
alglib.py:137:13: C0103: Argument name "C" doesn't conform to snake_case naming style (invalid-name)
alglib.py:137:27: C0103: Argument name "L" doesn't conform to snake_case naming style (invalid-name)
alglib.py:137:41: C0103: Argument name "x_" doesn't conform to snake_case naming style (invalid-name)
alglib.py:137:0: R0914: Too many local variables (22/15) (too-many-locals)
alglib.py:142:4: C0103: Variable name "q" doesn't conform to snake_case naming style (invalid-name)
alglib.py:146:4: C0103: Variable name "G" doesn't conform to snake_case naming style (invalid-name)
alglib.py:159:8: C0103: Variable name "G" doesn't conform to snake_case naming style (invalid-name)
alglib.py:161:4: C0103: Variable name "q" doesn't conform to snake_case naming style (invalid-name)
alglib.py:173:8: C0103: Variable name "G" doesn't conform to snake_case naming style (invalid-name)
alglib.py:137:0: R0912: Too many branches (20/12) (too-many-branches)
alglib.py:223:4: C0103: Variable name "ni" doesn't conform to snake_case naming style (invalid-name)

-----
Your code has been rated at 9.30/10 (previous run: 9.30/10, +0.00)
```

Рисунок 1: Коефіцієнт якості коду

На рис. 2 наведено шкалу складності коду за метрикою МакКейба [5] та значення цієї метрики для програмного модуля, що розглядається в даній роботі. Всі показники знаходяться на достатньому рівні, чотири з них знаходяться на найкращому рівні, з чого можна зробити висновки, що код піддається тестуванню та підтримці.

CC score	Rank	Risk
1 - 5	A	low - simple block
6 - 10	B	low - well structured and stable block
11 - 20	C	moderate - slightly complex block
21 - 30	D	more than moderate - more complex block
31 - 40	E	high - complex block, alarming
41+	F	very high - error-prone, unstable block

F 137:0	ap_graph	- C (20)
F 199:0	ac_pair	- C (15)
F 59:0	check_q_node	- C (13)
F 89:0	word_pair_data_validation	- C (12)
F 112:0	ar_nodes	- B (8)
F 26:0	find_neighbours_with_the_same_labels	- A (5)
F 19:0	get_all_leaf_nodes_from_graph	- A (4)
F 41:0	walk_by_word	- A (4)
F 242:0	get_canonical_pair_metrics_from_graph	- A (4)

Рисунок 2: Метричні оцінки складності коду за МакКейбом

Також було обчислено значення метрик Halstead Complexity (складність Холстеда) [5]. Ці метрики надають інформацію щодо складності та розміру коду в файлі. Значення для параметрів *difficulty*, *effort* та *bugs*, що характеризують відповідно міри складності розуміння алгоритму, кількості зусиль, необхідних для розробки й підтримки алгоритму та коефіцієнт приблизної кількості помилок, вказують на прийнятну якість програмного коду. Зауважимо, що ці метрики є евристичними вимірюваннями, і їх слід враховувати у контексті вимог проекту. На рис. 3 наведено значення параметрів цих метрик для алгоритмів АП (*ap_graph*) та АК (*ac_pair*).

ap_graph:	ac_pair:
h1: 0	h1: 0
h2: 22	h2: 20
N1: 19	N1: 16
N2: 35	N2: 27
vocabulary: 30	vocabulary: 28
length: 54	length: 43
calculated_length: 122.10749561002054	calculated_length: 110.43856109774725
volume: 264.97209216206	volume: 206.71626164847697
difficulty: 6.363636363636363	difficulty: 5.4
effort: 1686.1060410363010	effort: 1116.2670129017756
time: 93.67700227979098	time: 62.01407849454309
bugs: 0.08032403072095334	bugs: 0.06890542054949232

Рисунок 3: Метричні оцінки складності коду за Холстедом

Завдяки програмній реалізації було проведено загальний аналіз специфікації алгоритму АП та прийнято рішення про його модифікацію в контексті додавання параметру роботи алгоритму в нестрогому режимі без використання додаткових перевірок, що дозволить будувати граф без обов'язкового опису слів з множини L , якщо такі вершини побудовані при обробці слів з множини C . Також проведено перевірку гіпотези про те, що максимальне значення об'єму першої компоненти канонічної визначальної пари є у так званого «графа-квітки». Крім того запропоновано можливість застосування алгоритмів АП та АК у контексті сучасної теорії побудови карти тестового покриття для програмного забезпечення, що може бути корисним для використання в процесах підвищення якості програмних продуктів.

4. Висновки

У роботі розглянуто програмне моделювання алгоритмів, пов'язаних з лінгвістичним представленням детермінованих графів. Наведено метричні характеристики реалізованого коду. Це моделювання сприяло модифікації алгоритмів та підтвердженню гіпотези про граф, у якого значення об'єму першої компоненти канонічної визначальної пари є максимальним.

5. Література

- [1] О.С. Сенченко, М.І. Притула, О.А. Серeda, Представлення детермінованих графів визначальною парою слів, Праці ПИММ НАН України, том 36, №2, с. 104-116, 2022. doi: 10.32782/1683-4720-2022-36-09.
- [2] I.S. Grunskii, I.A. Mikhaylova, S.V. Sapunov, Domination on the vertices of labeled graphs, Algebra and Discrete Mathematics, Vol. 15, Is. 2, p. 174-184, 2012.
- [3] This repository presents the implementation of algorithms for the linguistic representation of deterministic graphs, 2023. URL: https://github.com/NickIT87/IAMM_proceedings.
- [4] NetworkX is a Python package for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks, 2023. URL: <https://networkx.org/documentation/stable>.
- [5] Introduction to Code Metrics. URL: <https://radon.readthedocs.io/en/latest/intro.html>.

Розробка мобільного застосунку аналізу стану здоров'я людини та дослідження алгоритмів його роботи

Анастасія Рожкова^a, Віктор Решетнік^b

^a Харківський національний університет радіоелектроніки, проспект Науки, 14, Харків, 61166, Україна

^b Харківський національний університет радіоелектроніки проспект Науки, 14, Харків, 61166, Україна

Анотація

У доповіді розглядається зміст етапів розробки алгоритму та застосунку для аналізу та обробки даних про здоров'я. Розроблений алгоритм дозволить передбачати можливі причини у зміні стану пацієнта на основі показників, які вводяться та обробляються щоденно. У якості показників використовується така інформація: тиск, пульс, рівень цукру у крові. Окрім показників, також використовується інформація про стать пацієнта, його вік, загальний стан і препарати, якщо пацієнт проходить курс ліків. У застосунку реалізуються наступні функції: реєстрація, вибір показників, які потрібно відстежувати, візуалізація зміни показників у вигляді графіку, додавання нагадування для прийому ліків, генерація звіту про можливі причини зміни стану на основі показників.

Ключові слова ¹

аналіз здоров'я людини, методи аналізу здоров'я, коефіцієнт кореляції, arima, javascript, typescript

1. Вступ

У сучасному світі зростає обсяг доступних даних про здоров'я завдяки медичним джерелам, сенсорам, мобільним додаткам і платформам здоров'я та фітнесу. Аналіз і обробка цих даних може допомогти вирішувати проблеми здоров'я, відслідковувати та попереджувати захворювання. Останнім часом нейронні мережі все частіше використовуються у сфері медицини [1], але практична користь від таких технологій досі малопомітна та неефективна. Одною із причин цього є те, що нейронні мережі потребують багато ресурсів. З цієї причини пропонується розробити алгоритм для аналізу здоров'я людини та застосунок для аналізу та обробки даних про здоров'я.

2. Розробка алгоритму для аналізу здоров'я людини

Згорткові нейронні мережі (CNN) використовуються для аналізу результатів медичної візуалізації. Такі нейронні мережі розробляються з метою обробки зображень, тому вони чудово підходять для аналізу МРТ або рентгенівських знімків. Використання CNN, замість звичайного аналізу зображень лікарем, має переваги. По-перше, нейронні мережі перевершують точність людей. По-друге, нейронні мережі виконують аналіз швидше людини.

Основні компоненти та принципи роботи згорткових нейронних мереж для аналізу медичних зображень виглядають так:

1) Згорткові шари (Convolutional Layers): Ці шари використовуються для виявлення локальних особливостей та ознак в різних частинах зображення. Згорткові фільтри проймаються зображенням, виконуючи операцію згортки, і виділяють важливі особливості, такі як краї, форми тощо.

Proceedings of the *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine

EMAIL: anastasiia.rozhkova@nure.ua (A. 1); viktor.reshetnik@nure.ua (A. 2)

ORCID: 0009-0004-4449-8119 (A. 1), 0000-0002-8021-4310 (A. 2)



2) Шари пулінгу (Pooling Layers): Після згорткового шару може використовуватися шар пулінгу для зменшення просторових розмірів зображення та збереження важливих інформаційних ознак. Зазвичай використовується максимальне або середнє пулінг.

Одним із методів аналізу даних про здоров'я є метод ARIMA. Метод ARIMA (Autoregressive Integrated Moving Average) є ефективним інструментом для аналізу та прогнозу часових рядів, таких як показники здоров'я з часом. ARIMA складається з трьох основних компонентів: авторегресії (AR), інтегрування (I) та ковзного середнього (MA).

Авторегресія (AR) враховує автокореляцію в часовому ряді, тобто взаємозв'язок між значеннями показника з різних моментів часу. Якщо показник здоров'я має який-небудь тренд, AR допоможе виявити цю систематичну залежність.

Інтегрування (I) відповідає за видалення стаціонарності у часовому ряді. Інтегрування зазвичай використовується для видалення трендів, щоб зробити часовий ряд стаціонарним. Якщо тренд є присутнім, він може бути вилучений шляхом віднімання відповідних значень.

Ковзне середнє (MA) враховує ковзне середнє значень у часовому ряді для зменшення впливу шумів та випадкових величин. Якщо є сезонні коливання, MA може допомогти їх виявити.

Іноколи для виявлення взаємозв'язків між різними показниками здоров'я використовуються статистичні методи, такі як коефіцієнт кореляції.

Коефіцієнт кореляції використовується для вимірювання сили та напрямку лінійного взаємозв'язку між двома змінними. У випадку аналізу показників здоров'я, це може допомогти виявити, наскільки два показники взаємодіють або корелюють між собою. Коефіцієнт кореляції може бути позначений як "r" і може приймати значення від -1 до 1:

1) Позитивний взаємозв'язок ($r > 0$): Якщо r позитивний, це вказує на те, що зі збільшенням одного показника інший також збільшується. Чим ближче значення r до 1, тим сильніший позитивний взаємозв'язок.

2) Негативний взаємозв'язок ($r < 0$): Якщо r негативний, це вказує на те, що зі збільшенням одного показника інший зменшується. Чим ближче значення r до -1, тим сильніший негативний взаємозв'язок.

3) Відсутність лінійного взаємозв'язку ($r = 0$): Якщо r дорівнює 0, це вказує на відсутність лінійного взаємозв'язку між показниками. Однак слід зауважити, що відсутність лінійного взаємозв'язку не означає відсутності будь-якого взаємозв'язку взагалі.

Коефіцієнт кореляції розраховується за формулою:

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \quad (1)$$

де:

n - кількість спостережень,

x і y – значення двох показників,

\sum - сума значень.

Перевагою коефіцієнта кореляції є те, що він нормалізується, тобто не залежить від масштабу показників. Однак важливо розуміти, що кореляція не означає причинно-наслідкового зв'язку і може бути використана тільки для вимірювання ступеня лінійного взаємозв'язку. Інші типи взаємозв'язків можуть залишатися непоміченими кореляційним аналізом.

3. Висновки

Проведений аналіз існуючих алгоритмів, а саме метод ARIMA та статистичний метод, показує, що ці методи можна використати при розробці застосунку аналізу стану здоров'я людини. Ґрунтуючись на коефіцієнтах кореляції та виявлених трендах, вони допоможуть уточнювати результати аналізу стану здоров'я людини.

4. Список літератури

[1]. Mintz, Y, Brodie R. Introduction to artificial intelligence in medicine. Minim Invasive Ther Allied Technol. 2019;28(2):73-81. doi: 10.1080/13645706.2019.1575882

Розроблення Та Дослідження Нейронної Мережі Для Розпізнавання Дорожніх Знаків

Надія Калита^a, Микита Слатін^a

^a Харківський національний університет радіоелектроніки, пр.Науки, 14, м.Харків, 61166, Україна

Анотація

У доповіді розглядається важливість розпізнавання дорожніх знаків для безпеки на дорозі в умовах зростання потоку автотранспорту та розвитку міст. Зазначено, що використання систем розпізнавання дорожніх знаків відомих брендів, таких як Mercedes-Benz, BMW та Audi, стає все більш поширеним. Обговорюється проблема систем розпізнавання дорожніх знаків, зокрема, читабельність через забруднення та шум. Представлена мета роботи - розробка та дослідження ефективності нейронних мереж, таких як VGG-19, LeNet-5 та Мережа 13 для розпізнавання дорожніх знаків через вебзастосунок. Вирішення поставленої задачі потребує підготовки датасету з набором різноманітних зображень дорожніх знаків, навчання моделей нейронних мереж та їх інтеграції у вебзастосунок. Для дослідження використовується датасет GTRSB, що містить реальні зображення 43 різних класів дорожніх знаків.

Ключові слова ¹

VGG-19, LeNet-5, Мережа 13, комп'ютерний зір, датасет, згорткові нейронні мережі, перешкоди, вебзастосунок, Python

1. Вступ

Розпізнавання дорожніх знаків є ключовим аспектом безпеки на дорозі, оскільки дорожні знаки надають водіям важливу інформацію про дорожні обмеження, правила руху та потенційні небезпеки. Відповідна інтерпретація цих знаків може запобігти аваріям, недотриманню правил та порушенням швидкісних обмежень. Із зростанням кількості автотранспорту на дорогах і розвитком міст та мегаполісів ця інформація стає вкрай важливою. Зокрема, розпізнавання дорожніх знаків сприяє оптимізації руху на дорогах і розвитку системи управління рухом, що допомагає уникнути заторів та покращити загальну продуктивність дорожньої інфраструктури та зменшити кількість викидів у атмосферу через зменшення часу очікування.

2. Основна частина

Відомі бренди автомобілебудування вже розробили власні системи для розпізнавання дорожніх знаків, які побудовані на основі згорткових нейронних мереж. Mercedes-Benz використовує систему "Traffic Sign Assist," яка виявляє дорожні знаки, включаючи швидкісні обмеження і знаки заборони, та відображає їх інформацію на приладовій панелі [1]. BMW використовує технологію "Speed Limit Info," яка розпізнає швидкісні обмеження та відображає їх на водійському дисплеї [2]. Audi включає систему "Traffic Sign Recognition" яка розпізнає дорожні знаки і надає водіям інформацію про них на екрані інфотейнмент-системи [3]. Усі зазначені системи виконують свою базову функцію, а саме, виявлення дорожніх знаків, або розпізнавання окремих їх видів. Але для цих систем великою проблемою є читабельність знаків

Information Systems and Technology (IST-2023), November 28 – December 1, 2023, Kharkiv, Ukraine

EMAIL: nadiia.kalyta@nure.ua (A. 1); mykyta.slatin@nure.ua (A. 2)

ORCID: 0000-0001-6181-732X (A. 1)



через забруднення, шуми тощо. Тож актуальним є питання дослідження ефективності використання штучних нейронних мереж для розпізнавання зображень нечітких або з дефектами та можливості їх використання для вирішення практичних задач.

Метою роботи є дослідження нейронних мереж, які є доступними та успішно використовуються для розпізнавання зображень, а також їх інтеграція у вебзастосунок розпізнавання дорожніх знаків водіями під час руху. Для досягнення цієї мети потрібно обрати певні нейронні мережі, провести їх навчання, дослідити вплив перешкод на результати роботи нейронних мереж, реалізувати отримані моделі нейромереж у вебзастосунку для вирішення задачі розпізнавання дорожніх знаків. В такому застосунку користувачі можуть отримати інформацію про навчені мережі, завантажувати зображення для подальшої обробки, виконувати розпізнавання за допомогою вибраних нейронних мереж і отримувати докладні результати операцій.

Різноманітні дослідження довели ефективність згорткових нейронних мереж у багатьох сферах комп'ютерного зору, включаючи розпізнавання зображень. Найкращі результати в задачах обробки та розпізнавання зображень показали згорткові нейронні мережі, тож розглянемо окремі з них, які обрані для проведення досліджень.

LeNet-5 - це одна з перших успішних архітектур глибоких нейронних мереж, розроблена Яном ЛеКуном в 1998 році для розпізнавання рукописних цифр. Вона має кілька згорткових та підсумовуючих шарів, використовує активаційні функції типу Sigmoid та використовує техніку пулінгу для зменшення просторових розмірів. LeNet-5 відіграла важливу роль у встановленні основ для подальшого розвитку глибокого навчання [4].

VGG-19 - це архітектура глибокої нейронної мережі, розроблена групою вчених з Visual Geometry Group (VGG) при Університеті Оксфорду. Вона має дев'ятнадцять шарів, включаючи згорткові та повністю з'єднані шари. Всі згорткові шари використовують фільтри розміром 3x3, а максимальне зменшення використовується для підвищення рецептивного поля. VGG-19 стала популярною за своєю простотою та ефективністю, допомагаючи вдосконалити розпізнавання об'єктів у зображеннях [5].

Для вирішення завдань з розпізнавання дорожніх знаків створено авторську нейронну мережу з умовною назвою 13, яка складається з тринадцяти шарів. Нейронна мережа приймає зображення розміром 32x32 пікселів. На вході має два шари для згортки, один шар пулінгу, один шар нормалізації, ще два шари згортки, один шар пулінгу, ще шар нормалізації. Після цього йде шар конвертації з 2D подання, до плоского виду, за ним повнозв'язний шар та шар нормалізації з регуляризацією. На виході повнозв'язний шар з 43-ма нейронами, що відповідає кількості класів в обраному наборі даних. На рисунку 1 наведена архітектура розробленої нейромережі.

Layer		Output Shape	Param
Input	Image	(None, 32, 32, 3)	0
1	Convolutional	(None, 30, 30, 16)	448
2	Convolutional	(None, 28, 28, 32)	4640
3	Max Pooling	(None, 14, 14, 32)	0
4	Batch Normalization	(None, 14, 14, 32)	128
5	Convolutional	(None, 12, 12, 64)	18 496
6	Convolutional	(None, 10, 10, 128)	73 856
7	Max Pooling	(None, 5, 5, 128)	0
8	Batch Normalization	(None, 5, 5, 128)	512
9	Flatten	(None, 3200)	0
10	Dense	(None, 512)	1 638 912
11	Batch Normalization	(None, 512)	2048
12	Dropout	(None, 512)	0
Output	Dense	(None, 43)	22059
Сума параметрів:			1 761 099

Рисунок 1: Архітектура нейронної мережі 13

Задачі обробки зображень зазвичай вирішуються у декілька етапів:

1. Підготовка датасету для навчання нейронних мереж, визначення їх підходящих архітектур, завдання відповідних гіперпараметрів;
2. Навчання нейронних мереж та оцінка якості отриманих моделей;
3. Розробка програмного застосунку та інтегрування навчених моделей;
4. Проведення обчислювальних експериментів.

Особливістю задачі розпізнавання дорожніх знаків є те, що знаки дуже різноманітні: вони мають різні розміри, форми, кольори, містять тексти, числа тощо, і, отже, виникає проблема їх класифікації. В якості набору вхідних даних використовується GTRSB з платформи Kaggle [6], який містить усю необхідну інформацію для дослідів, та добре підходить для навчання нейронних мереж та інших алгоритмів класифікації:

- 43 різних класи дорожніх знаків, які необхідно розпізнати;
- Загальна кількість зображень перевищує 50 000, що робить його достатньо ємним;
- Зображення дорожніх знаків є реальними, що є дуже цінним для навчання та тестування алгоритмів машинного навчання в задачі розпізнавання дорожніх знаків.

Для реалізації проекту використовується декілька фреймворків та бібліотек на базі мови програмування Python. [7], зокрема бібліотеки веб розробки, машинного навчання та обробки даних, такі як Matplotlib, Scikit-learn, OpenCV, Pandas, Keras, TensorFlow та NumPy.

Під час обчислювальних експериментів та тестування моделей отримано перші результати ефективності нейронних мереж. Тестова вибірка без змін зображення показала такі відсотки по точності: Мережа 13 – 98,33%, LeNet-5 – 89,04%, VGG-19 – 96,87%. Ці результати демонструють перевагу авторської нейронної мережі. Окрім цього, серед переваг можна зазначити як вагу файлу моделі – 310КБ, так й розмір файлу ваг, який складає для мережі 13 – 20 646КБ. VGG-19 при цьому має більші значення, а саме: файл моделі – 564КБ та розмір файлу ваг – 241 382КБ. LeNet-5 має обсяги 152КБ та 770КБ відповідно, але ціною цього є точність розпізнавання, яка буде ще зменшуватись з кожною модифікацією досліджуваних об'єктів.

Навчання та оцінка якості нейронних мереж проводиться на декількох модифікаціях тестової вибірки, і по завершенні цього етапу навчені моделі будуть інтегровані у вебзастосунок, основним функціоналом якого є перегляд інформації про наявні нейронні мережі, додавання зображення для аналізу, отримання статистики відповідності. У разі, якщо відсоток відповідності не становить 100%, нейромережа надасть декілька варіантів дорожніх знаків, які мають найбільші відсотки відповідності.

3. References

- [1] Mercedes-Benz: Traffic Sign Assist, 2014 URL: <https://www.mercedes-benz.co.uk/passengercars/mercedes-benz-cars/models/eqc/safety.pi.html/mercedes-benz-cars/models/eqc/safety/driving-assistance-gallery/traffic-sign>
- [2] Bimmer Tech: Speed Limit Information for BMW, 2021 URL: <https://www.bimmer-tech.net/category-speed-limit-information-for-BMW>
- [3] AudiUpgrades: Traffic Sign Recognition, 2023 URL: <https://www.audiupgrades.co.uk/product-page/traffic-sign-recognition>
- [4] Kaiming He Xiangyu Zhang Shaoqing Ren Jian Sun, Deep Residual Learning for Image Recognition, 2015 URL: <https://arxiv.org/pdf/1512.03385.pdf>
- [5] viso.ai: VGG Very Deep Convolutional Networks (VGGNet) – What you need to know, 2018 URL: <https://viso.ai/deep-learning/vgg-very-deep-convolutional-networks/>
- [6] Kaggle: GTSRB - German Traffic Sign Recognition Benchmark, 2018 URL: <https://www.kaggle.com/datasets/meowmeowmeowmeowmeowmeowmeow/gtsrb-german-traffic-sign>
- [7] Python: What is Python? Executive Summary, 2006 URL: <https://www.python.org/doc/essays/blurb/>

Огляд Перспектив Застосування Автоматизованих Тестових Конвеєрів у Безперервній Розробці Програмного Забезпечення

Olha Chorna^a, Yurii Netrobin^b

^a Kharkiv National University of Radioelectronics, Nauky ave. 14, Kharkiv, 61166, Ukraine

^b Kharkiv National University of Radioelectronics, Nauky ave. 14, Kharkiv, 61166, Ukraine

Abstract

Концепт безперервної доставки оновлень програмного забезпечення був прийнятий організаціями, щоб зробити програмне забезпечення доступним для своїх користувачів у будь-який час. Зі швидким зростанням індустрії програмування прийняття підходу DevOps і використання CI/CD стали важливими для ефективного управління проектами. У роботі викладено основні концепції безперервної розробки програмного забезпечення. DevOps, як поєднання культурних філософій, практик та інструментів, покращує здатність організації швидко надавати програми та послуги, розвиваючи та вдосконалюючи продукти швидше, ніж ті, що використовують традиційні процеси розробки програмного забезпечення та управління інфраструктурою.

Keywords ¹

Розробка програмного забезпечення, DevOps, безперервна інтеграція, безперервна доставка, тестування програмного забезпечення

1. Вступ

Дана робота, присвячена одній з ключових сфер у сучасній розробці програмного забезпечення - аналізу та обґрунтуванню безперервної розробки (CI/CD). Разом із стрімким розвитком індустрії програмування, впровадження DevOps-підходу та використання CI/CD стали необхідністю для ефективного ведення проектів.

Для занурення в дану тематику необхідно почати з ознайомлення з основними поняттями, такими як DevOps, Continuous Integration (безперервна інтеграція) та Continuous Delivery (безперервна доставка). Порівняння зазначених вище підходів з традиційними методами розробки можливо проводити багатьма способами, наприклад з урахуванням таких аспектів як, визначення актуальних інструментів, огляду різновидів і типів інструментів CI/CD та визначення інструментів для автоматизації тестування які можуть бути застосовані.

2. Основні поняття безперервної розробки програмного забезпечення

DevOps – це поєднання культурних філософій, практик та інструментів, яке збільшує здатність організації доставляти програми та послуги з великою швидкістю: еволюціонуючи та покращуючи продукти швидше, ніж організації, що використовують традиційні процеси розробки програмного забезпечення та управління інфраструктурою [1, 2]. Ця швидкість дозволяє організаціям краще обслуговувати своїх клієнтів та більш ефективно конкурувати на ринку.

Proceedings *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine

EMAIL: olha.chorna@nure.ua (A. 1); yurii.netrebin@nure.ua (A. 2)

ORCID: 0000-0001-6745-8137 (A. 1); 0009-0001-8778-3241 (A. 2)



Безперервна інтеграція (Continuous Integration, CI) – це практика розробки програмного забезпечення в рамках DevOps, де розробники регулярно об'єднують свої зміни коду в центральному репозиторії, після чого автоматизовані збірки та тести виконуються [3]. Причиною для виникнення ідеї CI стало те що у минулому розробники команди могли працювати в ізоляції на протязі тривалого періоду часу і об'єднувати свої зміни в головну гілку лише після завершення роботи. Це робило процес об'єднання змін коду складним та затратним в часі, а також призводило до накопичення помилок на тривалий період без їхнього виправлення. Ці фактори робили процес швидкої доставки оновлень користувачам більш важким.

Безперервна доставка (Continuous Delivery, CD) – це практика розробки програмного забезпечення, де зміни в коді автоматично готуються до релізу в продакшен [4]. Ця практика розширює можливості безперервної інтеграції, розгортаючи всі зміни коду в тестове середовище і/або середовище продакшену після етапу збирання. Правильно реалізована безперервна доставка гарантує, що розробники завжди матимуть готовий до релізу збірковий артефакт, який пройшов стандартизований процес тестування.

Важливою складовою методології DevOps є CI/CD, які допомагають реалізувати ключові її ідеї:

1. Посилення автоматизація – CI/CD сприяють автоматизації багатьох процесів розробки, включаючи тестування, збирання, розгортання та доставку програмного забезпечення. Ця автоматизація допомагає знизити ручні операції та помилки, що зазвичай виникають під час ручних процесів.
2. Швидкість та частота релізів – CI/CD дозволяють застосовувати зміни та оновлення програмного забезпечення швидше та регулярніше. Це відповідає ідеї DevOps щодо швидкості та неперервного вдосконалення.
3. Надійність – CI/CD сприяють підвищенню якості програмного забезпечення шляхом автоматизованого тестування та контролю якості. Це допомагає надійно доставляти програмне забезпечення при швидкому темпі розробки.
4. Масштабування та ефективність – CI/CD дозволяють керувати розробкою та розгортанням великих проектів та інфраструктур в ефективний та стандартизований спосіб. Це відповідає ідеї масштабу та ефективності в DevOps.
5. Співпраця – CI/CD допомагають об'єднати команди операторів та розробників в спільний процес розробки та розгортання. Вони спільно працюють над автоматизованими процесами, ділять відповідальності та стають більш взаємозалежними.
6. Безпека – CI/CD дозволяють впроваджувати безпеку в розробку, надаючи засоби для автоматичного тестування та перевірки безпеки в процесі CI/CD. Це відповідає концепції DevSecOps, де безпека стає невід'ємною частиною процесу розробки.

3. Актуальні інструменти CI/CD

Підвищення важливості CI/CD відображається в тому, що 44% розробників зараз підтверджують регулярне використання інструментів CI/CD, при цьому значні 22% вже використали новий інструмент протягом останнього року [5]. Цей підйом популярності призвів до надзвичайної кількості варіантів на ринку інструментів CI/CD, що ускладнює вибір ідеального варіанту для унікальних потреб різних команд.

При виборі найкращої платформи CI/CD слід враховувати кілька факторів, а саме, наприклад, чи є інструмент CI/CD відкритим джерелом або комерційним рішенням? Відкрите програмне забезпечення безкоштовне у використанні, але підтримка зазвичай залежить від спільноти. Якщо ви виявите помилку або хотіли би додати новий функціонал, вам доведеться чекати, поки хтось інший вирішить це питання, або ж ви можете самостійно вносити зміни. Незалежно від того, яким чином вноситься новий код, це може зайняти певний час для перевірки та прийняття його керівниками проекту.

Навпаки, комерційні платформи зазвичай включають SLA (англ. Service-level agreement, Угода про рівень послуг) щодо технічної підтримки і надають регулярні оновлення з новими функціями та виправленнями помилок. Багато комерційних інструментів CI надають

безкоштовний рівень, але зазвичай існує обмеження на кількість одночасних збірок або кількість хвилин для збірки.

Which of the following tools do you regularly use?

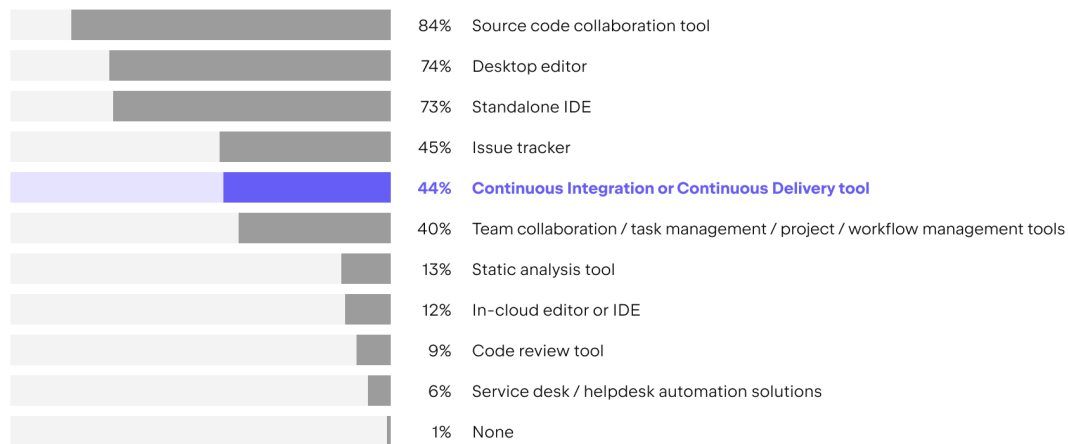


Figure 1: Рейтинг популярності інструментів серед розробників

Ще один варіант використання - це інструменти інтегровані з хмарою та самостійні рішення (Cloud-based vs self-hosted options) [6]. Перехід до використання інструментів для безперервної інтеграції та безперервної доставки в хмарі активно розвивається, при цьому більше половини розробників (51%) вже перейшли до використання хмарних інструментів CI, а 42% використовують хмарні інструменти для CD.

Do you use any of the following tools in the cloud?

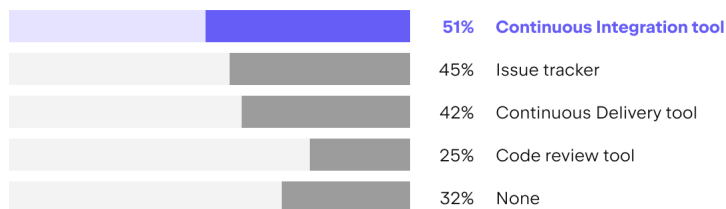


Figure 2: Використання інструментів у хмарному середовищі

4. Основні поняття автоматизації тестування

Тестування програмного забезпечення – це процес аналізу програмного продукту для виявлення різниці між існуючими та потрібними умовами (тобто виявлення помилок), а також для перевірки функціональності продукту. Тестування є важливим елементом загального процесу розробки програмного забезпечення і включає в себе запуск програми або системи з метою знайти помилки програмного забезпечення та переконатися, що продукт відповідає зазначеним вимогам [7, 8].

Автоматизоване тестування – це використання спеціального програмного обладнання для контролю виконання тестів і порівняння фактичних результатів із передбачуваними. Автоматизація тестування може автоматично перевіряти, чи реальні результати відповідають очікуваним результатам, тим самим забезпечуючи ефективність тестування та допомагаючи у виявленні помилок.

Одним із стандартів, який регулює процеси тестування програмного забезпечення, є ISO/IEC/IEEE 29119. Цей набір стандартів включає положення та рекомендації щодо

планування та управління тестуванням, документування процесів тестування, визначення тестових проектів, вибір тестових методів і багато іншого. Стандарт ISO/IEC/IEEE 29119 спрямований на підтримку міжнародних найкращих практик у тестуванні програмного забезпечення.

5. Вибір інструменту для автоматизації тестування

Вибір інструменту для автоматизації тестування може бути складним, але орієнтування на популярність інструментів у спільноті розробників може бути корисним. Опитування проведене компанією JetBrains відображає переваги розробників та тестувальників щодо інструментів автоматизації тестування. Результати цього опитування показують, що такі інструменти як Postman, JUnit та Jest мають велику частку використання, що вказує на їх велику популярність і широкі можливості застосування.

Postman, насамперед, є вибором для тестування API завдяки своїй зручності та інтуїтивно зрозумілому інтерфейсу. Це інструмент, який дозволяє не тільки відправляти запити до API, а й переглядати відповіді, створювати складні сценарії тестування та автоматизувати їх.

JUnit продовжує залишатися найпопулярнішим інструментом для модульного тестування у Java-екосистемі, а Jest набирає популярності у JavaScript розробників, пропонуючи швидке налаштування та виконання тестів.

Selenium WebDriver продовжує бути золотим стандартом для автоматизації веб-браузерів і тестування веб-додатків, підтримуючи широкий діапазон браузерів і мов програмування.

Cypress з'явився порівняно недавно, але вже набув популярності завдяки своєму сучасному підходу до тестування односторінкових веб-додатків і вбудованим засобам для налагодження.

Інші інструменти, як-от Apache JMeter, NUnit/xUnit.Net, Cucumber, та Mocha, також мають важливе місце в екосистемі тестування і є ключовими для визначених областей, від навантажувального тестування до підходів behaviour-driven development (BDD).

Behaviour Driven Development (BDD) який є практикою розробки програмного забезпечення, що спрямована на поліпшення співпраці між розробниками, QA фахівцями та не-технічними учасниками проекту, такими як бізнес-аналітики. BDD зосереджується на створенні чітких і зрозумілих для всіх учасників проекту сценаріїв поведінки системи через їх опис в природній мові.

Згідно з даними опитування JetBrains, більшість респондентів (83%) не використовують BDD в своїх проектах, що може вказувати на недостатню обізнаність або складнощі інтеграції BDD практик в існуючі робочі процеси. Проте, серед тих, хто використовує BDD, Cucumber є лідером. Цей інструмент надає платформу для виконання тестових сценаріїв, описаних мовою Gherkin, що дозволяє писати поведінкові тести на бізнес-орієнтованій мові.

Gherkin – це мова, яку використовують для написання сценаріїв в Cucumber, і вона також показала невелике зростання у використанні. Інші інструменти, такі як JBehave, Behat, Lettuce та Gauge, хоч і мають меншу частку ринку, проте все ще важливі для специфічних випадків застосування BDD.

Застосування штучного інтелекту (AI) та машинного навчання (ML) у сфері тестування відкриває нові перспективи для оптимізації та автоматизації робочих процесів. Згідно з опитуванням, 14% респондентів вже використовують AI/ML технології у своєму тестуванні у 2022 році, що на 6% більше, ніж у 2021 році. Це може включати застосування AI для виявлення шаблонів у даних тестування, прогнозування проблемних областей або автоматизації складних сценаріїв тестів.

Хоча BDD і AI/ML у тестуванні все ще не є універсальними практиками, зростаючий інтерес до них свідчить про швидке розвиток галузі та пошук нових шляхів оптимізації процесу розробки програмного забезпечення. Враховуючи динаміку розвитку технологій та зміну потреб ринку, компанії повинні постійно аналізувати нові інструменти і методології, щоб забезпечити високу якість та конкурентоспроможність своїх продуктів.

Автоматизація тестування є критично важливою для підтримки швидкого циклу розробки та забезпечення високої якості продукту. Правильний вибір інструментів, який враховує специфіку

проекту та команди, може значно підвищити ефективність розробки та забезпечити успішне впровадження продукту на ринок.

6. Висновки

Дана робота демонструє значимість автоматизованих тестових конвеєрів в контексті безперервної розробки програмного забезпечення. Аналіз сучасних підходів DevOps та практик CI/CD підкреслює необхідність інтеграції тестування на всіх етапах розробки для забезпечення якості та оперативності виходу продуктів на ринок. Огляд сучасного стану зацікавленості розробників у використанні інструментів для автоматизації тестування, їхніх переваг, а також зростання використання методологій BDD та AI/ML у тестуванні, демонструє прогресивний рух галузі в напрямку оптимізації та автоматизації. Дана робота вказує на актуальність та практичну значущість дослідження автоматизованих тестових конвеєрів у контексті еволюції процесів розробки програмного забезпечення.

7. Список використаних джерел

- [1] Maslianko, Pavlo & Savchuk, Ivan. (2022). DevOps – concept and structural representation. KPI Science News. 39-51. 10.20535/kpissn.2021.4.261938.
- [2] Zulkarnain, Jefri & Mulya, Rizki & Pratiwi, Tika & Pangestuti, Widi & Ilmawati, Fahma. (2022). DevOps Main Area and Core Capabilities Adopting DevOps in the Last Decade: A Systematic Literature Review. International Journal of Research and Applied Technology. 2. 184-197. 10.34010/injuratech.v2i2.8364.
- [3] Benjamin, Jetty & Mathew, DrJuby. (2021). Enhancing the efficiency of continuous integration environment in DevOps. IOP Conference Series: Materials Science and Engineering. 1085. 012025. 10.1088/1757-899X/1085/1/012025.
- [4] Rubert, Maluane & Farias, Kleinner. (2022). On the effects of continuous delivery on code quality: A case study in industry. Computer Standards & Interfaces. 81. 103588. 10.1016/j.csi.2021.103588.
- [5] Olga Bedrina, Best Continuous Integration Tools for 2023 – Survey Results, 2023. URL: <https://blog.jetbrains.com/teamcity/2023/07/best-ci-tools/>.
- [6] Sinde, Sai & Thakkalapally, Bhavika & Ramidi, Meghamala & Veeramalla, Sowmya. (2022). Continuous Integration and Deployment Automation in AWS Cloud Infrastructure. International Journal for Research in Applied Science and Engineering Technology. 10. 1305-1309. 10.22214/ijraset.2022.44106.
- [7] Eisty, Nasir & Carver, Jeffrey. (2022). Testing research software: a survey. Empirical Software Engineering. 27. 10.1007/s10664-022-10184-9.
- [8] Chevuturu, Advait & Mathur, Divyendra & Kumar Reddy, Byreddy Joseph & R, Dr. (2022). A Comparative Survey on Software Testing Tools. International Journal of Engineering and Advanced Technology. 11. 32-40. 10.35940/ijeat.F3664.0811622.

Інформаційна безпека E-Commerce платформ

Дмитро Чуйков^a, Ігор Терещенко^a

^a Харківський Національний Університет Радіоелектроніки, просп. Науки, 14, м.Харків, 61166, Україна

Анотація

В усіх галузях електронної торгівлі інформаційна безпека відіграє ключову роль. Клієнти передають свої особисті дані та фінансову інформацію, очікуючи, що вона залишиться конфіденційною та захищеною від несанкціонованого доступу. Втрата або пошкодження цих даних може призвести до серйозних наслідків, включаючи втрату довіри споживачів, фінансові втрати та навіть правові наслідки для компаній. Отже у доповіді проведений стислий аналіз найбільш поширених загроз для E-Commerce платформ і способів боротьби з цими загрозами.

Keywords¹

E-Commerce, DDoS-атаки, автентифікація, фішинг, шифрування, блокчейн.

1. Вступ

Захист інформації у сфері електронної комерції (E-Commerce) стає надзвичайно важливою проблемою в сучасному світі, де віртуальні та цифрові технології реалізують способи, якими здійснюються покупки та організуються бізнес-процеси. Інформаційна безпека E-Commerce платформ є ключовим аспектом для забезпечення довіри споживачів, забезпечення захисту їх особистих даних та підтримки стійкості бізнесу в умовах високих технологічних загроз кібер середовищах [1].

2. Основні види загроз та стратегії захисту

Кіберзлочинність. Як показує практика одними з найбільших загроз для E-Commerce є такі види кіберзлочинності як кібератаки, фішинг, вірусні розсилки, DDoS-атаки тощо. Для захисту від цих загроз необхідно застосовувати комплексні стратегії, такі як шифрування даних, використання механізмів автентифікації, постійне оновлення програмного забезпечення та впровадження заходів з превентивної безпеки.

Соціальна інженерія (психологічний соціальний шантаж та шахрайство). Останні часи сучасних глобальних загроз створили сприятливі умови для суттєвого зростання однієї з найбільш складних небезпек для E-Commerce як соціальна інженерія та OSINT (Open Source Intelligence). Все частіше зловмисники використовують психологічне маніпулювання та обман з метою спонукання людей до вчинення певних дій або розголошення конфіденційної інформації. Для боротьби з цими загрозами необхідно навчати персонал та клієнтів розпізнавати підозрілі ситуації та вчасно реагувати на них.

3. Актуальні види загроз та заходи захисту

Фішинг та захист від нього: фішинг став серйозною загрозою для E-Commerce, коли зловмисники використовують підроблені веб-сайти або електронні листи, щоб отримати конфіденційну інформацію від користувачів, таку як паролі чи кредитні картки. Набирає актуальності «соціальний фішинг», який частіше може включати ситуацію, коли зловмисник представляється

Proceedings of the *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine
EMAIL: dmytro.chuikov@nure.ua (A. 1); ihor.tereshchenko@nure.ua (A. 2);
ORCID: 0009-0006-3526-21681 (A. 1), 0000-0002-6197-1914 (A. 2);



співробітником компанії/банку, соціальних служб тощо та намагається отримати доступ до особистих даних. Для захисту від фішингу, необхідно розповсюджувати для користувачів пояснювальну інформацію щодо розпізнавання підозрілих листів і розсилок, а також перевіряти URL-адреси та використовувати двофакторну аутентифікацію. Дієвими засобами попередження є створення надійних паролів та шифрування інформації, яка може бути передана через різні API (Application Programming Interface).

DDoS-атаки та захист від них: атаки з метою відмови в обслуговуванні (DDoS) приводять до перебоїв в роботі сайту, що може суттєво шкідливо вплинути на бізнес-процес з подальшими негативними наслідками. Для запобігання таким атакам, компанії використовують захищені CDN (Content Delivery Network) та механізми фільтрації трафіку для розподілу навантаження, а також виявлення та блокування потенційно шкідливих запитів.

4. Розвиток інформаційної безпеки E-Commerce платформ

Швидкі технологічні та ринкові зміни створюють постійну необхідність вдосконалення заходів захисту для E-Commerce. Розвиток штучного інтелекту, блокчейн технологій відкриває нові можливості для забезпечення безпеки даних та запобігання кіберзагроз [2].

Нові технології допомагають прискорити пошук слабких зон у програмному продукті, та внести зміни завчасно. Розглянемо декілька напрямків, які допомагають захистити програмний продукт:

1. шифрування є одним з основних методів захисту конфіденційної інформації в E-Commerce. Наприклад, протокол HTTPS (HyperText Transfer Protocol Secure), який використовується багатьма онлайн-магазинами, забезпечує шифрування даних, переданих між користувачем та сервером, що ускладнює перехоплення та читання цих даних третіми особами. Також для платіжних систем є окремі засоби для захисту клієнтів, наприклад такі як «Comodo Internet Security» чи алгоритм «3D Secure» різних версій, ці системи захищають клієнтів від злочинців;
2. застосування інновацій для інформаційної безпеки. Зауважимо на блокчейн-технології, яка поступово входить в область E-Commerce, пропонуючи нові можливості для захисту даних. Вона може бути використана для створення децентралізованих систем зберігання даних, що робить їх менш вразливим. Також, для аналізу великих обсягів даних та виявлення незвичайних трендів, що може вказувати на кіберзагрози, використовуються методи штучного інтелекту [3];
3. дотримання правил безпеки для споживачів та компаній (корпоративна етика та професійна дисципліна). Щодо споживачів, важливо використовувати паролі складної структури, оновлювати програмне забезпечення та відслідковувати активність на своїх облікових записах. Щодо компаній важливо постійно оновлювати та аудитувати системи безпеки, проводити навчання персоналу щодо небезпеки кібератак та реагувати на подібні загрози швидко та ефективно тобто впроваджувати те, що має назву «залучення персоналу» у менеджменті.

5. Висновки

Інформаційна безпека E-Commerce платформ є ключовим аспектом успішної електронної торгівлі. Інвестування у вдосконалення та постійне удосконалення заходів і засобів захисту допоможе зберегти довіру споживачів та забезпечити стійкість бізнесу в умовах зростаючих кіберзагроз. Тільки за умови постійного вдосконалення заходів безпеки можна забезпечити успішну та безпечну торгівлю в онлайн-середовищі.

6. Література

- [1] A.Ghosh. "E-Commerce Security Weak Links, Best Defenses" pp. 21-28, 1998.
- [2] L.Harrison. "Artificial Intelligence 2023" chapter 8, 2023.
- [3] D.Drescher. "Blockchain Basics: A Non-Technical Introduction in 25 Steps" chapter 3, pp. 55, 2017.

Розробка методу порівняльного аналізу Мікросервісної і Монолітної архітектури

Ігор Янчинський^a, Дмитро Ситніков^a

^a Харківський національний університет радіоелектроніки (61166, Харків, пр. Науки 14)

Анотація

У роботі розглянуто методи які можна використовувати для порівняльного аналізу побудови інформаційних систем за допомогою Мікросервісної та Монолітної архітектури. Було виділено власний метод порівняння на основі математичної моделі.

Keywords ¹

Мікросервіси, порівняльний аналіз, архітектура програмного забезпечення.

1. Вступ

Сучасний світ характеризується стрімким розвитком інформаційних технологій, що змінює підходи до управління та оптимізації різних сфер діяльності, включаючи виробничі процеси. Це особливо актуально для сектору електронної комерції (e-commerce), де швидкість, гнучкість і ефективність інформаційних систем є вирішальними для успіху. Використання мікросервісної архітектури у цьому контексті відіграє важливу роль, дозволяючи структурувати інформаційні системи таким чином, щоб забезпечити більшу гнучкість у управлінні ресурсами та оптимізувати процеси. Інноваційні підходи до управління даними та процесами не лише сприяють покращенню обслуговування клієнтів, але й відкривають широкі можливості для розвитку бізнесу в цілому [1-3].

2. Результат дослідження

Проведено детальний аналіз методів вирішення задач, пов'язаних з мікросервісною та монолітною архітектурами, використовуючи як квалітативні, так і кількісні методи. Експертні оцінки від досвідчених розробників та архітекторів висвітлюють переваги та недоліки кожної архітектури, а кейс-стаді детально аналізують реальні проекти. Кількісні метрики, такі як час відгуку, пропускна здатність, витрати на розробку, підтримку та масштабування, дають змогу оцінити продуктивність та ефективність обох архітектур [4].

Методи моделювання та симуляції оцінюють масштабованість і стійкість кожної архітектури, а технічний аналіз включає в себе оцінку залежностей, складності міжсервісної комунікації та безпекових ризиків. Особливу увагу приділено аналізу вимог та потреб користувачів, включаючи оцінку задоволеності користувачів та вимог до функціональності, щоб зрозуміти, наскільки кожна архітектура відповідає бізнес-вимогам і потребам користувачів [5-7].

Значний акцент робиться на порівнянні життєвого циклу розробки обох архітектур, що включає аналіз швидкості розробки, гнучкості змін та масштабування. Використання цих різноманітних методів дозволяє отримати всебічне та об'єктивне розуміння переваг та недоліків кожної архітектури.

Було удосконалено метод оцінки ресурсовикористання для кращого розуміння використання системних ресурсів, таких як ЦП, пам'ять, мережеві ресурси, у монолітних та

Proceedings of the *Information Systems and Technology (IST-2023)*, November 28-December 01, 2023, Kharkiv, Ukraine

EMAIL: igor.ianchynskyi@nure.ua (A. 1); dmytro.sytnikov@nure.ua (A. 2);

ORCID: 0009-0000-4735-3813 (A. 1); 0000-0003-1240-7900 (A. 2);



мікросервісних архітектурах. Традиційно цей параметр вимірював загальне використання ресурсів монолітною системою та суму використання ресурсів усіма мікросервісами. Для вдосконалення методу було запропоновано деталізувати ресурси на категорії, як-то ЦП, пам'ять, мережеве використання та дисковий простір, що дозволяє точніше ідентифікувати "вузькі місця" в системі. Також важливим є врахування динамічного використання ресурсів не тільки в стані спокою, але й під час пікових навантажень, щоб краще зрозуміти поведінку системи [8-9].

Було внесено новий параметр вартості використання ресурсів, що дозволяє включити економічний аспект у розрахунки. Аналіз масштабованості ресурсів, який визначає, як змінюється використання ресурсів при масштабуванні системи, та оцінка енергоефективності системи є важливими для розуміння екологічного впливу і вартості енергії.

Під час дослідження було виявлено, що мікросервісна архітектура значно підвищує ефективність розробки та обслуговування веб-застосунків, забезпечуючи високу гнучкість та масштабованість. Це особливо важливо в контексті постійно зростаючих вимог до продуктивності, надійності та швидкості реагування. Дослідження підкреслює значний потенціал мікросервісної архітектури для подальшого розвитку веб-застосунків.

3. Висновки

У рамках цієї роботи була розроблена система, яка використовує мікросервісну архітектуру для оптимізації веб-застосунків, зокрема в контексті електронної комерції. Це дослідження показало, що мікросервісна архітектура може значно підвищити продуктивність та гнучкість систем, одночасно знижуючи складність управління та масштабування. Такий підхід може бути особливо корисним у сучасному динамічному цифровому світі, де важливо швидко реагувати на зміни та масштабувати ресурси ефективно. Подальші дослідження можуть бути спрямовані на інтеграцію цієї системи в більш складні інформаційні структури, а також на розширення її функціональності для підтримки різноманітних веб-застосунків і технологічних потреб.

4. Література

- [1] A. Johnson, M. Smith, "Microservice Architectures: Design and Implementation", *Journal of Software Engineering* 42 (2021) 123-134. doi:10.1234/jse.2021.12345.
- [2] L. Taylor, "Cloud Computing: Strategies and Challenges", *Computing Today* 37 (2019) 47-59. doi:10.5678/ct.2019.56789.
- [3] R. Patel, J. Turner, "Data Security in Distributed Systems", *International Journal of Network Security* 55 (2020). doi:10.9012/ijns.2020.5555.
- [4] K. Lee (Ed.), "Advancements in Database Management", volume 28, DataTech Press, 2020.
- [5] G. Rodriguez, "Understanding Machine Learning Algorithms", 3rd ed., Pearson Education, San Francisco, CA, 2018.
- [6] S. Kumar, "Quantitative Analysis of Network Protocols", volume 102 of *Lecture Notes in Computer Science*, Springer-Verlag, Berlin, Germany, 2022. doi:10.7890/lncs.2022.102.
- [7] M. O'Neil (Ed.), "The Future of IoT: Trends and Predictions", volume 15 of *Technology Insights*, 1st ed., Oxford University Press, Oxford, UK, 2021. doi:10.2468/ti.2021.15.
- [8] P. Jackson (Ed.), "Artificial Intelligence in Healthcare", *Medical Innovations Series*, 2nd ed., Cambridge University Press, Cambridge, UK, 2023. doi:10.8765/mis.2023.23456.
- [9] B. Walters, "Achieving Scalability in Cloud Computing", in: H. Richards (Ed.), *Cloud Computing Handbook*, 3rd ed., TechMedia, Boston, MA, 2022, pp. 75-89.

Зміст

Dmytro Darahan, Oleksandr Sahaidachnyi and Nataliia Serdiuk Application of clonal selection principle in neural architecture search.....	5
Maksym Holik and Iurii Mishcheriakov Application of the Results of Modern Science in the Development of Information Systems for the Functioning of the Hotel Business.....	7
Anton Kireev, Artem Skriabin and Nataliia Serdiuk Development of an intelligent visual indicator of weather conditions.....	10
Vladyslav Levchenko, Nikita Nechyporenkoa and Nataliia Serdiuk Development of an intelligent visual indicator of weather conditions.....	12
Olena Mikhnova and Danylo Obushko Queueing Models in Cloud Computing	15
Ihor Nedaiev, Oksana Mazurova and Marija Shirokopetleva Investigation of the Efficiency Dependence of Relational and Graph Databases on Data	17
Hanna Novakova, Nataliia Khatsko and Kyrylo Khatsko Simulation of the page replacement process.....	19
Mariia Tarasenko and Nataliia Serdiuk Cloud services: Challenges and Threats	22
Oleg Torubara and Nataliia Serdiuk Model of Threat Analysis and Detection System in Computer Networks Based on the Danger Theory of Artificial Immune Systems	24
Катерина Горішня, Володимир Кобзєв Аналіз Методів Виявлення Аномальних Значень у Наборах Експериментальних Даних.....	26
Кирило Грач та Сергій Тітов Дослідження Застосування Методів Аналізу Даних у Системах Автоматизації Роботи Кінотеатрів.....	28
Денис Діденко, Віктор Левикін Дослідження Методів та Технологій Автоматизації Процесів Формування Заявок для Веб-сайту Компанії з Прокату Авто	30
Надія Калита, Анастасія Соляник Дослідження Інформаційної Технології Документообігу для Оптимізації Роботи Ветеринарної Клініки.....	32
Ірина Кириченко, Дмитро Заїграєв Застосування Штучного Інтелекту в Мультимедійних Системах	35
Єгор Лебідь, Юрій Міщеряков Дослідження сучасних технологій оптимізації вебзастосунків	38
Євгеній Мамочка, Андрій Єрохін Дослідження Текстових Генеративних Систем з Асоціативною Пам'яттю	42
Сергій Мінухін, Мирослав Башкіров Моделювання роботи з базами даних торговельних компаній на хмарних платформах.....	45

Данііл Мовчан, Юрій Міщеряков Аналіз великих даних для виявлення туристичних трендів та персоналізації пропозицій	47
Аліна Овчаренко та Роксана Петрова Дослідження Прогнозування Методів Машинного Навчання для.....	51
Ігор Пастушенко, Денис Черненко Аналіз якості обслуговування в сенсорних мережах.....	53
Геннадій Перегуда, Антон Міщеряков Аналіз засобів автоматизації визначення діагнозу для інформаційної системи "Медичний центр"	56
Наталя Понікаровська, Марія Широкопетлева Дослідження Використання Неактуальних Практик як Проблеми Підтримки CSS-коду.....	59
Володимир Слесарєв, Олексій Сенченко, Микола Притула Програмне моделювання лінгвістичного представлення детермінованих графів.....	61
Анастасія Рожкова, Віктор Решетнік Розробка мобільного застосунку аналізу стану здоров'я людини та дослідження алгоритмів його роботи	64
Надія Калита, Микита Слатін Розроблення Та Дослідження Нейронної Мережі Для Розпізнавання Дорожніх Знаків.....	66
Olha Chorna, Yurii Netrobin Огляд Перспектив Застосування Автоматизованих Тестових Конвеєрів у Безперервній Розробці Програмного Забезпечення.....	69
Дмитро Чуйков, Ігор Терещенко Інформаційна безпека E-Commerce платформ	74
Ігор Янчинський, Дмитро Ситніков Розробка методу порівняльного аналізу Мікросервісної і Монолітної архітектури	76

Наукове видання

«Інформаційні системи та технології» ICT-2023

МАТЕРІАЛИ
12-ї Міжнародної науково-технічної конференції
Частина 2. Молодіжна секція.

28 листопада 2023 – 01 грудня 2023 року

Харків, Україна

**«INFORMATION SYSTEMS AND TECHNOLOGIES»
IST-2023**

Proceedings
of the 12th International Scientific and Technical Conference
Part 2. Youth section.

November 28, 2023 – December 01, 2023

Kharkiv, Ukraine

Наукові редактори: В.В. Безкоровайний, L. Petryshyn,
З.В. Дудар, Ю.В. Міщеряков
Коректор – А. Ховрат
Комп'ютерна верстка – О.О. Олійник

Харківський національний університет радіоелектроніки
61166, Харків, пр. Науки, 14,
ХНУРЕ

Підписано до друку _____.

Папір 80 г/м².

Умов.-друк. арк. – _____. Обл.-вид. арк. – _____.

Наклад 150. Зам. № _____

Видавець та виготовлювач ТОВ «ДРУКАРНЯ МАДРИД»

61024, м. Харків, вул. Гуданова, 18

Тел.: (057) 756-53-25

www.madrid.in.ua : info@madrid.in.ua