МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЧЕРКАСЬКИЙ ДЕРЖАВНИЙ ТЕХНОЛОГІЧНИЙ УНІВЕРСИТЕТ

ВІСНИК
ЧЕРКАСЬКОГО ДЕРЖАВНОГО ТЕХНОЛОГІЧНОГО УНІВЕРСИТЕТУ

Науковий збірник

Том 29,
№ 3. 2024

ЧЕРКАСИ
2024

ISSN: 2306-4412
E-ISSN: 2708-6070

Засновник:
Черкаський державний технологічний університет

Рік заснування: 1996

Рекомендовано до друку та поширення
через мережу Інтернет Вченою радою
Черкаського державного технологічного університету
(протокол № 3 від 21 жовтня, 2024 р.)

Державна реєстрація: Ідентифікатор медіа R30-04613.
Рішення Національної ради України з питань телебачення і радіомовлення
№ 1916, протокол № 17 (30.05.2024 р.).

Науковий збірник входить до переліку наукових фахових видань України
Категорія «Б». Спеціальності:
122 – Комп'ютерні науки, 123 – Комп'ютерна інженерія, 125 – Кібербезпека,
126 – Інформаційні системи та технології, 151 – Автоматизація та комп'ютерно-інтегровані технології, 172 – Телекомунікації та радіотехніка
(наказ Міністерства освіти і науки України від 28 грудня 2019 року № 1643)

Науковий збірник представлено у міжнародних наукометричних базах даних, репозитаріях та пошукових системах: Bielefeld Academic Search Engine (BASE), Crossref, CiteFactor, Ulrich's Periodicals Directory, WorldCat, Eurazian Scientific Journal Index (ESJI), Directory of Open Access Journals (DOAJ), Open Ukrainian Citation Index (OUCI), Наукова періодика України, Національна бібліотека України імені В. І. Вернадського (НБУВ), Dimensions, UCSB Library, Google Академія, German Union Catalogue of Serials (ZDB), Бібліотека Університету Осло, Бібліотека Університету Халла, SOLO - Search Oxford Libraries Online, Інститут Європейського університету, Бібліотека Лейпцизького університету (UBL), Бібліотека Кембриджського університету

Адреса редакції:
Черкаський державний технологічний університет
18006, бульв. Шевченка, 460, м. Черкаси, Україна
E-mail: info@bulletin-chstu.com.ua
https://bulletin-chstu.com.ua/uk

MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE
CHERKASY STATE TECHNOLOGICAL UNIVERSITY

BULLETIN
OF CHERKASY STATE TECHNOLOGICAL UNIVERSITY

Scientific Collection

Volume 29,
No. 3. 2024

CHERKASY
2024

ISSN: 2306-4412
E-ISSN: 2708-6070

Founder:
Cherkasy State Technological University

Year of foundation: 1996

Recommended for printing and distribution
via the Internet by the Academic Council
of Cherkasy State Technological University
(Minutes No. 3 of October 21, 2024)

State Registration: Media identifier R30-04613.
Decision of the National Council of Television and Radio Broadcasting of Ukraine
No. 1916, Minutes No. 17, dated 30.05.2024.

The scientific collection is included in the list of Professional Scientific Publications of Ukraine
Category "B". Specialties: 0613 – Software and applications development and analysis,
0611 – Computer use, 0612 – Database and network design and administration,
0714 – Electronics and automation
(order of the Ministry of Education and Science of Ukraine No. 1643, dated December 28, 2019)

The scientific collection is presented international scientometric databases, repositories and scientific systems: Bielefeld Academic Search Engine (BASE), Crossref, CiteFactor, Ulrich's Periodicals Directory, WorldCat, Eurazian Scientific Journal Index (ESJI), Directory of Open Access Journals (DOAJ), Open Ukrainian Citation Index (OUCI), Scientific Periodicals of Ukraine, Vernadsky National Library of Ukraine (VNLU), Dimensions, UCSB Library, Google Scholar, German Union Catalogue of Serials (ZDB), University of Oslo Library, University of Hull Library, SOLO - Search Oxford Libraries Online, European University Institute, Leipzig University Library (UBL), Cambridge University Library

Editors office address:
Cherkasy State Technological University
18006, 460 Shevchenko Blvd., Cherkasy, Ukraine
E-mail: info@bulletin-chstu.com.ua
https://bulletin-chstu.com.ua/en

Редакційна колегія

Editorial Board

Зміст / Contents

BULLETIN of Cherkasy State Technological University

Journal homepage: https://bulletin-chstu.com.ua/en

Vol. 29 No. 3. 2024

UDC 004.491
DOI: 10.62660/bcstu/3.2024.10

Malware development: From early viruses to modern cyber threats

Denys Kovalchuk*
Postgraduate Student
International Humanitarian University
65009, 33 Fontanska Rd., Odesa, Ukraine
https://orcid.org/0009-0003-2302-8698

Abstract. Malware is one of the biggest threats in the digital environment, as it is constantly evolving and becoming more dangerous. The purpose of this study was to analyse the evolution of malicious software. Historical, comparative and empirical analysis and an assessment of existing security technologies were used to achieve this goal. The main findings revealed several key stages in the development of malicious software. Historical analysis has shown that the evolution of malware has gone through several significant stages, from simple viruses and worms to complex threats such as ransomware and spyware. These changes were driven by technological advances and increased attack capabilities, which allowed malicious software to use new vectors of influence and deception methods. A comparative analysis of modern cyber threats revealed the key characteristics and differences between different types of malwares, including their specific distribution methods and vulnerabilities. It was found that new threats have a more complex architecture and use more innovative tactics, which significantly complicates their detection and neutralisation. An empirical analysis involving the use of threat detection tools provided specific data on malware behaviour in action. A review and testing of modern security methods, including antivirus solutions, intrusion detection systems, and firewalls, showed their strengths and weaknesses, as well as their effectiveness in detecting and preventing new threats. The results of the study highlighted the need for continuous improvement of protection methods, which is critical for effective control of modern cyber threats

Keywords: evolution of information attacks; digital security; intrusion detection systems; security technologies; computer incident analysis

INTRODUCTION

In today's digital environment, cybersecurity is critical to protecting data and systems from threats and attacks. The theoretical foundations of cybersecurity include understanding the main types of cyber threats, such as malicious software, phishing, ransomware, and ongoing threats of cyber-attacks. Information security principles, including privacy, data integrity, and availability, as well as security techniques such as cryptography, intrusion detection systems, and multi-layer protection, are also important aspects. In the context of rapid technology development and an increase in the number of cyber-crimes, the key aspects are the adaptability of security systems and their ability to quickly respond to new threats. In particular, various organisations and institutions face new challenges, such as the growing complexity of malicious software, infrastructure attacks, and threats to personal data. Existing security systems often do not keep up with the rapid development of attacking methods, which creates significant gaps in information security. These challenges highlight the need for continuous improvement of protection methods and the development of new strategies to effectively counter modern cyber threats. In general, malicious software and attacks, such as ransomware and phishing, are

Article's History: Received: 25.05.2024; Revised 10.08.2024; Accepted 21.10.2024

Suggested Citation:
Kovalchuk, D. (2024). Malware development: From early viruses to modern cyber threats. *Bulletin of Cherkasy State Technological University*, 29(3), 10-20. doi: 10.62660/bcstu/3.2024.10.

*Corresponding author

becoming more complex and adaptive. N. Antonenko *et al.* (2022) investigated current malware detection methods, emphasising the need for new protection mechanisms, which shows the need to improve existing methods. O. Marchenko (2023) analysed the effectiveness of antivirus solutions for critical infrastructure and found that although modern solutions show high results, some security gaps still remain. This confirms the need for further improvements in the protection of critical systems. I. Zulkovska *et al.* (2021) found that none of the existing signature and heuristic methods are perfectly effective, which confirms the need to combine new approaches, such as machine learning and graphical visualisation, to improve the accuracy of threat detection. In turn, the study by I. Galuzin & G. Naiman (2021) showed that vulnerability management technologies, such as Qualys solutions, are effective, but require further improvement and integration with new technologies. In addition, A. Alchi *et al.* (2024) noted that neural networks can significantly improve the accuracy of malware detection by being able to detect patterns in large data sets, although there are problems with the need for significant amounts of training data. This points to gaps in existing methods that require further study. The results of the study by S. Kumar & G. Nagar (2024) emphasised the need for adapted threat models for less developed organisations, which is supported by an analysis of historical incidents and modern sophisticated attacks.

Moreover, M. Qumer & S. Ikrama (2022) examined how artificial intelligence can improve corporate security by detecting and responding to cyber threats. This shows the importance of new technologies in improving security. H.N. Durmuş Şenyapar (2024) emphasised the importance of regular software updates and employee training to improve cybersecurity in digital marketing. The findings of these studies highlight the need for a comprehensive approach to protection that includes both technological and human aspects. The study by P. Ravi *et al.* (2024) demonstrated a machine learning system for detecting ransomware and malware, which indicates the prospects for new approaches in this direction. In addition, R.J. Anderson (2020) analysed current cybersecurity challenges, such as attacks through mobile devices and cloud services, highlighting the need for innovative security solutions. These studies confirm the need for continuous improvement of methods of protection and adaptation to new threats.

Therefore, understanding the evolution of malware is important for improving the level of cybersecurity. Analysis of current research shows that cyber threats continue to evolve, which requires continuous improvement of protection methods and raising awareness of new threats. The purpose of the study was to analyse the evolution of malicious software, and the tasks included identifying the main stages of malware development and evaluating the effectiveness of modern cybersecurity methods.

MATERIALS AND METHODS

For a comprehensive understanding of the evolution of malicious software, a detailed historical analysis was conducted, including the study of its development from initial to modern threats. This analysis examined the first computer viruses and traced their transformation into more complex and diverse types of malicious software. Historical analysis revealed key stages in the development of malware, changes in attack methods and targets, and noted the impact of technological progress on the evolution of threats. This approach involved the analysis of similar scientific studies, which provided a clear picture of how malware technologies and techniques have changed over time.

Next, a comparative analysis of various types of modern cyber threats was carried out, such as viruses (Chernobyl, ILOVEYOU), worms (Blaster, Sasser), Trojans (Zeus, Emotet), rootkits (Stuxnet, Alureon), spyware (Pegasus, CoolWebSearch), botnets (Mirai, Kraken) and ransomware (WannaCry, CryptoLocker). The comparative analysis identified key characteristics and differences between different types of malwares, including their specific distribution methods and vulnerabilities that they use to achieve goals. As part of this analysis, the effectiveness of various methods for detecting and protecting against these threats was also investigated. Various tools and platforms, such as VirusTotal and Cuckoo Sandbox, were used for practical analysis of modern cyber threats. These tools allowed to get data about real-world examples of malicious software programmes, their behaviour, and distribution methods. VirusTotal provided code analysis using various antivirus engines, while Cuckoo Sandbox allowed running data in an isolated environment and investigating its behaviour. Empirical analysis helped to obtain specific data on how malicious software interacts with systems and what protection methods are most effective for detecting and neutralising them.

Evaluation of current malware protection methods included analysis of various technologies and approaches, such as antivirus programmes, intrusion detection systems, firewalls, and user education programmes. Each of these methods was evaluated in terms of its effectiveness in combating modern cyber threats. The study tested and compared antivirus solutions (CCleaner, Norton Antivirus, Bitdefender, McAfee, Avast, TotalAV, Intego, Malwarebytes, Panda), anti-intrusion systems (Snort, Suricata, Malwarebytes, Yara, Hybrid Analysis, ReversingLabs), and firewalls. The evaluation included checking functionality, usability, impact on system performance, and the effectiveness of detecting and preventing new threats. In addition, the role of user education in reducing the risks of cyber threats, including the effectiveness of trainings, was analysed. The results of this assessment helped to formulate recommendations for improving cybersecurity in the face of modern threats.

RESULTS

The evolution of malicious software is one of the most pressing issues in the modern digital world. From the first computer viruses to modern complex cyber threats, the development of malicious software requires constant analysis and the creation of effective protection strategies. In general, the types of malicious software include viruses, worms, trojans, rootkits, spyware, botnets, and ransomware (Fig. 1).

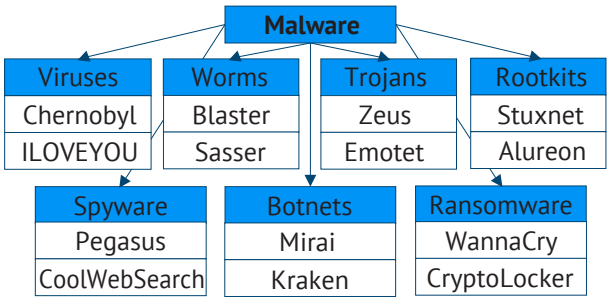

Figure 1. Types of malware and examples
Source: compiled by the author

Viruses are programmes that infect other files and systems by inserting their own code into them. Viruses can self-copy and spread to other computers through infected files. They can corrupt or delete data, slow down systems, or even destroy hardware components. Some viruses can change the functionality of programmes and systems. Worms are self-replicating programmes that are distributed over networks and the Internet, usually without the need for user interaction. They can cause network congestion, system crashes, or use infected devices to attack other systems. In turn, Trojans are programmes that disguise themselves as innocent or useful software, but actually perform malicious actions. They can give attackers access to the victim's system, steal sensitive information, or download other malicious software. Rootkits are a set of tools that allow attackers to gain administrative access to a computer and hide their activities. Rootkits can mask their presence, protect other malicious software from detection, and allow attackers to permanently access the system without the user's knowledge. As for spyware, they collect information about users without their knowledge or consent. Spyware can collect data about personal habits, online activity, passwords, financial information, and other sensitive data. Botnets are networks of infected computers that are centrally controlled to perform malicious tasks. They are used to launch distributed Denial of Service (DDoS) attacks, send spam, steal data, or perform other criminal activities. In addition, ransomware is malicious software that encrypts files on the victim's computer and demands a ransom for decrypting them. Ransomware blocks access to important files or systems and requires payment (usually in cryptocurrency) to restore access. The original malware, such as Creeper and Elk Cloner, appeared in the 1970s and 1980s and had a fairly simple architecture (Matthews, 2022). Creeper, created by Robert Thomas, was one of the first self-copying viruses, and its main purpose was to demonstrate the ability to transfer between computers (Fig. 2). Elk Cloner, created in 1982, was one of the first viruses to infect floppy disks and caused a demo message every 50th system startup.

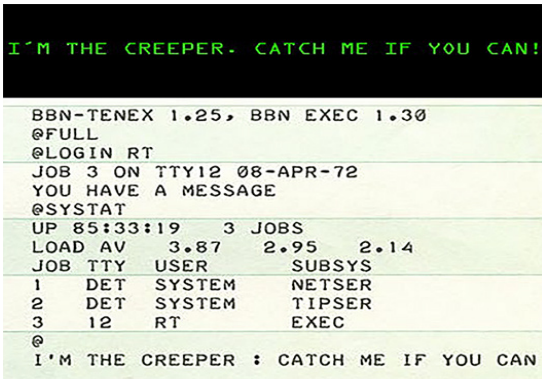

Figure 2. Creeper virus message displayed on infected computers and its source code
Source: compiled by the author

Since the early 1990s, viruses have become more complex. For example, the Chernobyl virus since 1998 included mechanisms that allowed it not only to destroy data, but also to infect the BIOS (Basic Input/Output System) of the computer, which made it difficult to restore the system (Hanna, 2021). Over time, viruses began to use polymorphism mechanisms that made it difficult for antivirus software to detect them by changing their code each time they were launched. Since the 2000s, new types of malicious software have emerged, including Trojans, worms, and rootkits. Blaster and Sasser, which became known in 2003, have demonstrated the ability to quickly spread across networks and exploit vulnerabilities in operating systems (OS) (Karlberg, 2004). These worms exploited vulnerabilities in Windows for their attack, which confirmed the need for regular security updates. Between 2010 and 2020, malicious software underwent significant changes, demonstrating new levels of complexity and technological development. At this time, such types of malwares as ransomware, botnets, and targeted attacks appeared. CryptoLocker, which appeared in 2013, was one of the first ransomware programmes that encrypted files on victims' computers and demanded a ransom in cryptocurrency. Botnets such as Mirai have exploited vulnerabilities in Internet of Things (IoT) devices to create large-scale DDoS attacks that lead to large-scale infrastructure failures. During this period, there was also an increase in targeted attacks on various organisations and state structures. After 2020,

malicious software has become even more complex and specialised. Ransomware attacks such as Conti and DarkSide continue to hit organisations, including critical infrastructure, and require huge amounts in the form of ransomware, often in cryptocurrency (Fig. 3). At the same time, spyware such as Pegasus continues to be used to collect information on mobile devices, in particular, due to vulnerabilities in iOS and Android (Nemchick, 2023). By attacking personal data and communications, these programmes endanger the privacy and security of users on a global level. Consequently, there are a large number of different malicious programmes that continue to evolve with each passing decade, becoming more dangerous, which underscores the need for continuous improvement of cybersecurity technologies.

Modern cyber threats have a significant impact on businesses, individuals, and government agencies. Ransomware causes serious damage by encrypting data and demanding ransom, which can lead to significant financial losses and failures in critical infrastructure. Spyware threatens the privacy of personal data by collecting information without users' knowledge. Botnets can exploit vulnerabilities in IoT to carry out DDoS attacks, leading to large-scale failures in network systems and services. Rootkits allow criminals to mask their presence in the system and collect information without users noticing. Thus, there are many different types of malwares that have significantly changed global security and cyber infrastructure (Table 1). Attacks on critical infrastructure demonstrate how cyber threats can turn into large-scale problems for society. The growing number of DDoS attacks on large service providers and network resources indicates how malicious software can create global disruptions to internet services.

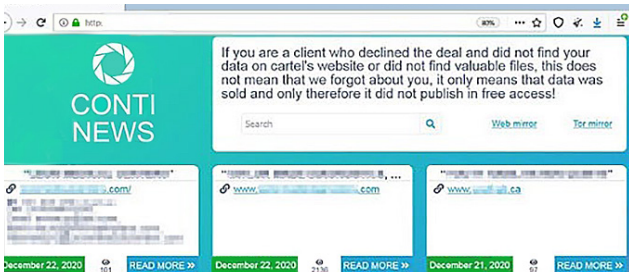

Figure 3. Conti website, where hackers publish stolen data, demanding ransom from victims
Source: compiled by the author based on J. Dalman & H. Smith (2021)

Table 1. Comparison of malware and cyber threats

Type of malware	Main features	Impact on systems
Viruses	Self-copying, infecting files	Data corruption or deletion
Worms	Fast distribution across networks	Network congestion, failures
Trojans	Covert interference, abuse of rights	Data leakage, system compromise
Rootkits	Masking presence, maintaining access	Long-term covert surveillance
Spyware	Collection of information without the user's knowledge	Breach of confidentiality
Ransomware programmes	File encryption, extortion	Loss of access to data, financial losses

Source: compiled by the author

For a better understanding, it is important to consider in detail exactly how malware functions and how they differ. Viruses are attached to files or uploaded to the boot sectors of disks. When an infected file is launched, the virus activates and starts copying its code to other files on disk. This process can spread to other computers if infected files are transferred over networks or specific media. Viruses can cause data corruption, system slowdowns, or even physical damage to hardware components. For example, Chernobyl is a virus that is activated on a specific day and can destroy data on the hard disk and in the computer's BIOS, which makes it difficult to restore the system. ILOVEYOU is a virus that spreads through email and causes numerous infections around the world. In turn, worms exploit vulnerabilities in network protocols or software for self-propagation. They do not require user interaction and often automatically scan networks for vulnerable devices. After finding the vulnerability, the worm infects the device and uses it for further spread. Examples include Blaster, a worm that exploits a Windows vulnerability to spread automatically and cause system crashes, and Sasser, a worm that infects systems through a Windows vulnerability and also causes computers to crash.

Trojans disguise themselves as legitimate or useful programmes that the user can download and install without suspicion. Once installed, the Trojan performs malicious actions, such as giving attackers access to the system, stealing sensitive data, or downloading additional malware. For example, Zeus is a Trojan that steals bank details by replacing web pages to collect login data. While Emotet is a Trojan that gains access to the system and downloads other malicious programmes, stealing data and causing violations in the system. Moreover, rootkits modify the OS or other

applications to hide their presence and activity. They can replace system files or change system settings, making them difficult to detect. Rootkits are often used for long-term monitoring or providing permanent access to the system without the user's knowledge. For example, Stuxnet is a rootkit specifically designed to attack nuclear facilities that modifies control systems to disrupt centrifuges. Alureon is a rootkit that hides other malware and steals data.

Spyware infiltrates the system and collects information about users without their knowledge. They can track online activity, collect input data, collect passwords, browser history, and other sensitive data. There is such a spyware programme as Pegasus, which exploits vulnerabilities in mobile systems to collect personal data and communications. And the CoolWebSearch programme changes the browser settings for collecting information and advertising. There are also botnets consisting of a network of infected devices that are controlled centrally through command servers. These devices can be used to coordinate attacks such as DDoS, sending spam, or data theft. For example, the Mirai botnet exploits vulnerabilities in IoT devices to create large DDoS attacks that cause failures in Internet resources. Kraken is a botnet used to attack websites and send spam.

Ultimately, ransomware encrypts files on the victim's computer and blocks access to them. Next, they demand a ransom to decrypt the data. They usually use strong encryption algorithms, which makes it difficult to recover files without a decryption key. One example is WannaCry, a ransomware that exploits vulnerabilities in Windows to encrypt data and demands a ransom in cryptocurrency. CryptoLocker also encrypts files and requires a ransom to recover them. It turns out that malicious programmes not only complicate conventional methods of protection, but also change approaches to cybersecurity. Attack detection systems such as Intrusion Detection System (IDS) and firewalls must constantly adapt to new forms of threats. The emergence of new types of malicious software requires improved security technologies and increased resources for monitoring and responding to incidents.

Malware monitoring and analysis are critical to ensuring cybersecurity. For this purpose, various tools and platforms were used to help identify, investigate, and respond to cyber threats. Among these tools were VirusTotal and Cuckoo Sandbox, which provide powerful malware analysis capabilities. VirusTotal is a free online service that allows checking files and Uniform Resource Locator (URL) for malicious code using dozens of antivirus programmes and other analysis tools (Fig. 4). When a file is uploaded to VirusTotal, it is checked by various antivirus systems that analyse the file for known malware samples. VirusTotal also provides information about potential threats based on an analysis of the file's behaviour and characteristics.

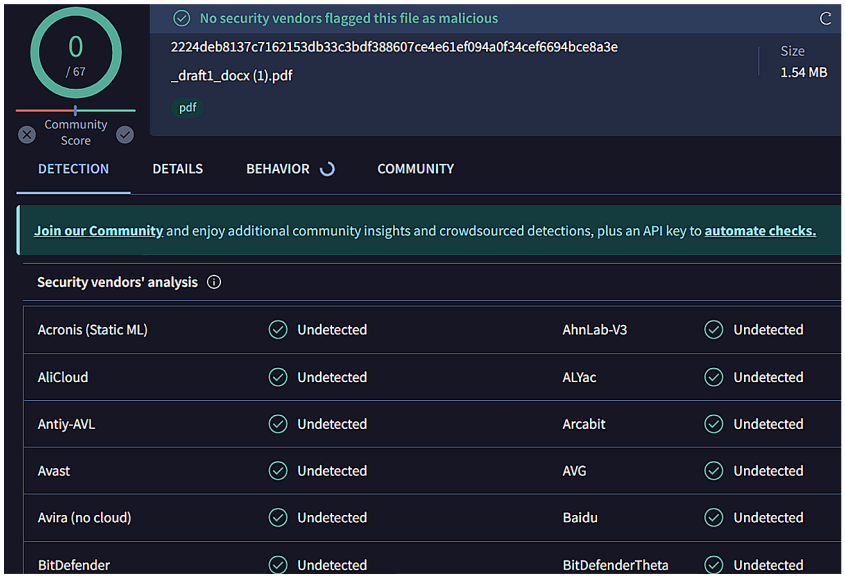

Figure 4. Example of using VirusTotal to check file security

Source: compiled by the author

The Figure 4 shows how the file is sent to VirusTotal, checked by various antivirus systems and analytical tools, and the result is provided to the user. The Cuckoo Sandbox platform is also deserves mentioning, as it is suitable for dynamic malware analysis (Fig. 5). It creates an isolated environment that simulates a real OS, in which malware can be executed to monitor its behaviour. Cuckoo Sandbox collects detailed information about the programme's actions, such as registry changes, file access, network activity, and creates reports that help to understand how the programme interacts with the system.

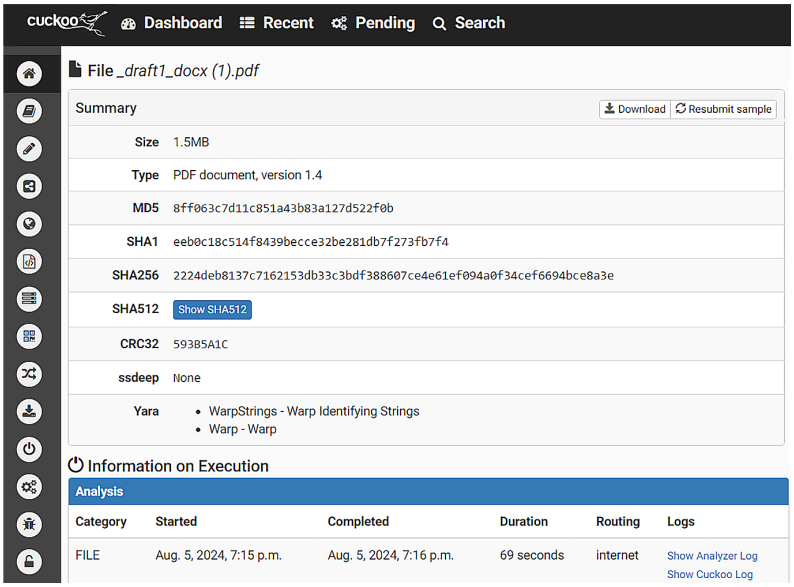

Figure 5. Example of using Cuckoo Sandbox to check file security
Source: compiled by the author

The Figure 5 illustrates how Cuckoo Sandbox creates an isolated environment for malware execution, tracks its behaviour, and collects data for further analysis. In addition to these tools, there are many other platforms that specialise in detecting malware. For example, Snort is an open-source intrusion detection system that monitors network traffic and analyses it for malicious activity. It detects abnormal behaviour, analyses packets, and monitors traffic. Suricata is a powerful open-source IDS/Intrusion Prevention System (IPS) that supports deep packet analysis and network monitoring. Its functions include checking network traffic, analysing threats, recognising protocols, and blocking attacks. There is also Malwarebytes, an antivirus and antispyware tool that specialises in detecting and removing malicious software and spyware, viruses, trojans, and adware. Yara is a tool for creating and using templates to detect and classify malicious files. It detects and analyses malicious patterns using templates and threat recognition. In turn, Hybrid Analysis is an online platform for dynamic and static malware analysis. It analyses file behaviour, collects information about requests, system changes, and network activity. ReversingLabs offers malware analysis platforms that include tools for static and dynamic analysis. This tool conducts research of malicious samples, detection of threats, and tracking the origin of malicious software. In addition, there are tools such as FireEye, OpenDXL, Threat Grid, SIEM systems, SOAR platforms, and EDR solutions. These tools and platforms play a key role in cybersecurity, helping organisations to effectively detect, analyse, and respond to malware and cyber threats.

In general, the development of malicious software encourages continuous progress in security systems and new cybersecurity technologies. With the growing complexity and diversity of cyber threats, cybersecurity professionals are forced to develop new solutions to detect and counter these threats. Modern security methods include innovative approaches and technologies that evolve in response to the emergence of new types of malwares. Antivirus programmes are among the oldest tools in the fight against malicious software. Modern antivirus systems use a combination of traditional signature techniques and more modern technologies, such as machine learning and behavioral analysis. Signature methods are based on recognising known malware patterns, while behavioural analysis detects new or unknown threats based on their activity in the system. These programmes include CCleaner (Fig. 6). CCleaner, although not an antivirus, is a popular tool for cleaning the system of temporary files, cache, and other unnecessary data. It helps to improve overall system performance and remove potentially unwanted applications. While CCleaner does not provide specific malware protection directly, using it can help to keep system clean and reduce the chance of infection.

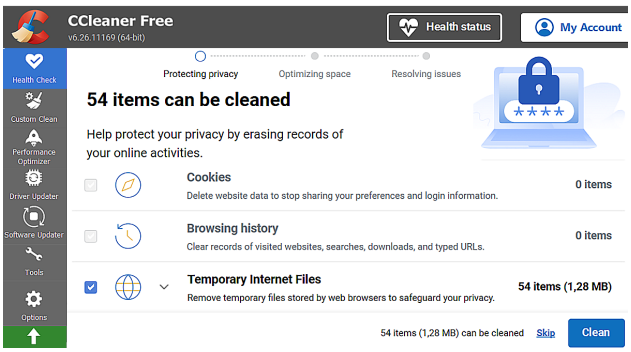

Figure 6. Example of using CCleaner utility software
Source: compiled by the author

Attention should also be paid to Norton Antivirus, which is one of the most well-known antivirus solutions. It provides a high level of protection due to powerful signature databases and behavioural analysis. Norton has effective features for detecting and blocking malicious software and spam mailings. However, its resource intensity can affect the speed of the system, and there are also constant updates that can be intrusive for users. The strong point of Bitdefender is its security technology and low impact on system performance. This application uses innovative security techniques such as machine learning and anti-phishing to ensure a high level of security. Bitdefender also offers features to protect sensitive data and ensure online activity. This solution is marked by a good balance between powerful protection and low resource consumption. In turn, McAfee offers comprehensive protection that includes not only antivirus features, but also protection against malicious websites and real-time threats. Its security mechanisms are robust, but sometimes McAfee can be noticeable in the system and consume a significant number of resources. In addition, Avast offers a free version with basic protection, including paid versions with advanced features. It performs well in virus detection and has a user-friendly interface. Avast provides privacy protection and blocking unwanted programmes. Thus, Norton Antivirus provides a high level of protection, but can be somewhat resource-intensive and have privacy issues. Bitdefender offers a balanced approach with powerful protection and low resource consumption. McAfee has comprehensive protection, but can consume a significant amount of system resources. Avast offers a user-friendly interface and a free version, but sometimes it can affect the system speed. And CCleaner is useful for maintaining overall system performance, but it does not replace specialised anti-malware solutions. However, Norton, TotalAV, McAfee, Intego, Bitdefender, Malwarebytes, and Panda are considered the best antivirus programmes of 2024 (Fig. 7).

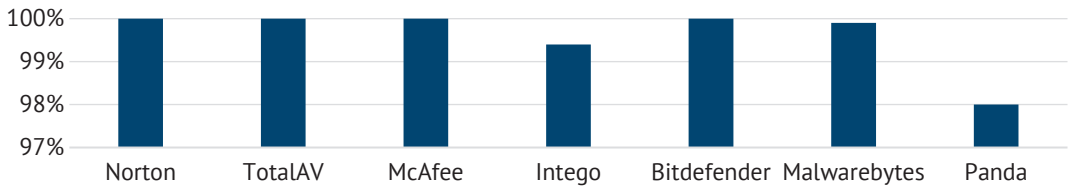

Figure 7. Rating of antivirus programmes in 2024

Source: compiled by the author

As for intrusion detection systems, they monitor network traffic and system activity to detect abnormal behaviour that may indicate an attack attempt (Fig. 8). IDS can be either signature-based or anomaly-based. Signature-based systems compare traffic to known attack patterns, while anomaly-based systems detect deviations from normal behaviour that may indicate new threats. There are systems such as OSSEC, Zeek, Security Onion.

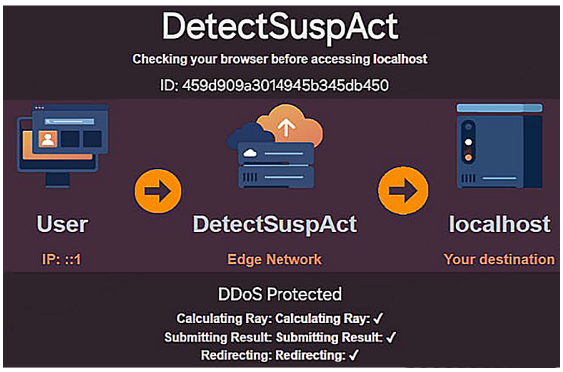

Figure 8. Example of using IDS

Source: A. Tsymbal (2024)

Firewalls are also important, which monitor incoming and outgoing network traffic based on a set of security rules and policies. Modern firewalls include deep packet analysis features that allow detecting malicious traffic and blocking it before it reaches end devices. They can be either hardware or software and provide multi-level protection. Examples include Cisco ASA, Palo Alto Networks Next-Generation Firewall, Check Point Firewall, and Sophos XG Firewall. An important aspect of cybersecurity is the education of users. Training users in the basics of cybersecurity, such as detecting phishing emails, and recognising malicious attachments, is critical to reducing the risk of malicious software infecting the system. Educational programmes often include security training and attack simulations to help users raise their awareness and readiness for cyber threats.

Regarding the effectiveness of modern protection methods, several key aspects have been identified. First, the combination of antivirus programmes, IDS, firewalls, and user education provides multi-level protection, which is important for reducing risks. Secondly, modern security systems are constantly being improved to deal with new types of malicious software. Machine learning and behaviour analysis technologies play a key role in identifying new threats. Thirdly, enabling proactive methods, such as regular software updates and system patches, helps to reduce vulnerabilities that can be exploited by attackers. Thus, anti-malware

methods are constantly evolving in response to new cyber threats. The integration of various approaches and technologies, together with user education, provides comprehensive protection against malware and reduces the risks of cyber threats.

DISCUSSION

The study showed the importance of implementing integrated security systems in modern networks, where the integration of the latest security technologies plays a key role in ensuring data integrity and confidentiality. Analysis of the findings in the context of existing studies indicates the importance of a systematic approach to the problem of cybersecurity, in particular, in the context of new threats and security technologies. The importance of such an analysis lies in the fact that it allows identifying new trends and approaches in the field of information security and supporting practical recommendations based on the results of previous research. This contributes to the further development and improvement of security methods, which is critical in the rapidly changing cyber environment.

The results of this study showed that the integration of the latest cryptographic algorithms and security technologies significantly improves the level of information protection. This correlates with the findings of W. Stallings & L. Brown (2018), who examined current cybersecurity principles, including cryptographic algorithms and general security principles. Their study confirmed the importance of implementing effective cryptography methods to ensure data security, which is consistent with the results obtained, which focuses on improving existing security methods through the latest cryptographic solutions. Meanwhile, M. Bishop (2018) addressed a wide range of issues related to modern security techniques, including attack analysis and intrusion detection. This study also confirmed that the integration of new security technologies increases the effectiveness of intrusion detection systems and reduces the risks that can arise from modern attacks. This highlights the importance of constantly updating and adapting security methods in the face of new threats, as shown in this paper, where the emphasis was placed on the latest methods of ensuring system security.

The study noted an increase in the complexity of modern attacks and malware, which requires advanced protection measures. In turn, the J. Ferdous *et al.* (2023) confirmed that the integration of multi-layer security systems can effectively counteract new types of threats, such as cryptojacking and attacks on IoT devices. In addition, this study is consistent with the conclusions of M. Jartelius (2020), who emphasised the importance of a proactive approach to threat detection to prevent data leaks, confirming the effectiveness of the integrated security systems implemented in this paper. The results of the study also indicate the importance of implementing innovative methods in the field of cybersecurity, such as automation and the use of the latest machine learning algorithms. This correlates with the findings of M. Conti *et al.* (2018), considered the security issues in IoT, and with the results of the study by Z. Chen (2020), where modern deep learning algorithms for malware protection were investigated. This study confirmed that the use of the latest technologies, such as deep learning, can significantly improve the effectiveness of protection systems against new types of cyber threats. In addition, the results of this study, which highlighted the effectiveness of post-quantum cryptography in data protection, are consistent with the study by P.N. Kokare *et al.* (2024), which compares traditional cryptography with post-quantum algorithms. Also important is the contribution of I. Legárd (2020), which focused on raising staff awareness of information security, which is critical to maintaining a high level of protection. This study has confirmed that an effective awareness-raising programme can have a significant impact on reducing the risks of cyber threats.

The results showed that the effectiveness of data protection can be improved by using the latest machine learning algorithms and automated systems. This is consistent with the conclusions of N. Akhter *et al.* (2021), who described how hackers exploited vulnerability in SolarWinds Orion software suite to compromise systems. They emphasised that timely detection and elimination of such vulnerabilities is crucial to prevent large-scale data leaks. In addition, the results of this study, which emphasised the importance of integrating multi-layer security systems to prevent attacks, correlate with the study by T. Olaniyan (2021), who also described the hacking of FireEye through the SolarWinds Orion update platform. The analysis showed that an attack on such a platform can lead to significant consequences for the security of organisations, which emphasises the need for comprehensive protection and rapid response to threats. Moreover, the results of this study showed the importance of continuously implementing the latest machine learning techniques in cybersecurity. T. Khatun (2024) also reviewed malware development and strategies in detail, which is consistent with the results of this study, which focused on the evolution of malware and the need to improve detection methods. This highlights the importance of continuous monitoring and adaptation to new threats.

It is worth noting that the results of this study focus on the evolution of malware and methods of combating it, which has common aspects with the papers by S. Okhanashvili (2023). The paper highlighted the importance of an integrated approach to improving cybersecurity, considering technical, programme, and organisational aspects. This study also confirms the importance of a comprehensive approach in the fight against cyber threats, but focuses on the evolution of malicious software and changes in its complexity over time. In addition, the results obtained share common features with the studies by K.K. Dewangan *et al.* (2024) and N.S. Janoti *et al.* (2024). In the first case,

the researchers investigated cyber threats to intelligent transport systems, focusing on new technologies such as blockchain and 5G to improve security. The current study also explored innovative approaches to security, but with an emphasis on general trends in the evolution of malicious software. In the second case, the researchers considered challenges in Cyber Threat Intelligence and the role of data analysis in protecting against cyber threats. The results also focused on the importance of data to combat evolutionary threats, but focused more on changes in the malware itself and its impact on cybersecurity. Ultimately, the study by O.C. Obi *et al.* (2024) also shares similarities with this study, as it offered a comprehensive overview of the current cybersecurity landscape, focusing on various threats such as malware and advanced persistent threats. However, the current study confirmed this trend, in particular in the context of the evolution of malware. Overall, the conducted study confirms the findings of other researchers on the evolution of malware and highlights the importance of comprehensive strategies to combat modern cyber threats. This is important for developing more effective strategies to protect against cyber threats in the future.

CONCLUSIONS

A study of the evolution of malicious software has confirmed that the development of malicious software has gone through several key stages, from simple viruses to modern complex cyber threats, such as ransomware and spyware. This study highlighted how changes in attack technologies and tactics lead to a constant increase in complexity and potential malware damage. Analysis of modern threats has shown that malicious software actively uses advanced technologies, including artificial intelligence, to bypass security systems. This creates new cybersecurity challenges that require organisations to implement more comprehensive and dynamic security strategies. Regular software updates, the introduction of intrusion detection systems, and the continuous improvement of cybersecurity tools are critical to countering modern threats. This study also highlighted the growing use of new attack methods, such as phishing and business email hacking. These new methods require organisations to implement additional security measures, such as multi-factor authentication and regular employee training. The use of advanced tools to analyse and monitor network traffic is also critical for early detection of suspicious activity. The importance of collaboration between different organisations and cybersecurity professionals is a key to improving security. Sharing of information about new threats and vulnerabilities, joint development of security methods and educational programmes can significantly improve the effectiveness of cybersecurity.

The main recommendations for improving cybersecurity include the introduction of regular updates of systems and software, the use of modern detection and prevention tools, the development of comprehensive protection strategies covering technical and organisational measures, and active cooperation with other experts in the field of cybersecurity. However, the study has some limitations, including a limited number of specific cases for analysing the evolution of malware and the complexity of assessing the impact of the latest technologies on all aspects of cybersecurity. Further research may include a more detailed analysis of new types of attacks, emerging technologies, and an assessment of their impact on various industries and sectors.

ACKNOWLEDGEMENTS

None.

CONFLICT OF INTEREST

None.

REFERENCES

[1] Akhter, N., Aziz, O., & Hussain, T. (2021). Latest trends in the cybersecurity after the solar wind hacking attack. *Foundation University Journal of Engineering and Applied Sciences*, 1(2). doi: 10.33897/fujeas.v1i2.347.

[2] Alchi, A., Dodiya, K., & Niveditha, V.S. (2024). Impact of neural network on malware detection. In K. Kaushik & I. Sharma (Eds.), *Next-generation cybersecurity* (pp. 219-241). Singapore: Springer. doi: 10.1007/978-981-97-1249-6_10.

[3] Anderson, R.J. (2020). *Security engineering: A guide to building dependable distributed systems*. London: Wiley. doi: 10.1002/9781119644682.

[4] Antonenko, N., Dihtyar, Ya., & Krykun, N. (2022). Modern methods of fighting computer viruses. *Economy and Society*, 43. doi: 10.32782/2524-0072/2022-43-51.

[5] Bishop, M. (2018). *Computer security: Art and science*. Boston: Addison-Wesley.

[6] Chen, Z. (2020). Deep learning for cybersecurity: A review. In *International conference on computing and data science* (pp. 7-18). Stanford: IEEE. doi: 10.1109/CDS49703.2020.00009.

[7] Conti, M., Dehghantanha, A., Franke, K., & Watson, S. (2018). Internet of Things security and forensics: Challenges and opportunities. *Future Generation Computer Systems*, 78(2), 544-546. doi: 10.1016/j.future.2017.07.060.

[8] Dalman, J., & Smith, H. (2021). *Under attack: Protecting against Conti, DarkSide, REvil and other ransomware*. Retrieved from https://www.crowdstrike.com/blog/how-to-defend-against-conti-darkside-revil-and-other-ransomware/.

[9] Dewangan, K.K., Panda, V., Ojha, S., Shahapure, A., & Jahagirdar, S.R. (2024). Cyber threats and its mitigation to intelligent transportation system. In *Symposium on international automotive technology*. Warrendale, Pennsylvania: SAE International. doi: 10.4271/2024-26-0184.
[10] Durmuş Şenyapar, H.N. (2024). Digital marketing in the age of cyber threats: A comprehensive guide to cybersecurity practices. *The Journal of Social Science*, 8(15), 1-10. doi: 10.30520/tjsosci.1412062.
[11] Ferdous, J., Islam, R., Mahboubi, A., & Islam, M.Z. (2023). A review of state-of-the-art malware attack trends and defense mechanisms. *IEEE Access*, 11, 121118-121141. doi: 10.1109/ACCESS.2023.3328351.
[12] Galuzin, I., & Naiman, G. (2021). Vulnerability management of corporate information systems based on QUALYS solutions. *Modern Information Security*, 46(2), 26-31. doi: 10.31673/2409-7292.2021.020708.
[13] Hanna, K.T. (2021). *Chernobyl virus*. Retrieved from https://www.techtarget.com/searchsecurity/definition/Chernobyl-virus.
[14] Janoti, N.S., Rohan, Rida, & Negi, N. (2024). Strategic perspectives on cyber threat intelligence: A comprehensive analysis. *International Journal for Research in Applied Science and Engineering Technology*, 12(4), 524-529. doi: 10.22214/ijraset.2024.59816.
[15] Jartelius, M. (2020). The 2020 data breach investigations report – a CSO's perspective. *Network Security*, 2020(7). doi: 10.1016/S1353-4858(20)30079-9.
[16] Karlberg, L.A. (2004). *Sasser faster than Blaster*. Retrieved from https://www.nyteknik.se/nyheter/sasser-snabbare-an-blaster/425179.
[17] Khatun, T. (2024). *Malware – unmasking the pervasive cyber threat of 2023*. Retrieved from https://www.researchgate.net/publication/382442937_Malware-_Unmasking_the_Pervasive_Cyber_Threat_of_2023.
[18] Kokare, P.N., Vora, D., Patil, S., Kotecha, K., Khairnar, V., Choudhury, T., & Kulkarni, A. (2024). *Post quantum cryptography: A survey of past and future*. Retrieved from https://www.researchgate.net/publication/382398375_Post_Quantum_Cryptography_A_survey_of_Past_and_Future.
[19] Kumar, S., & Nagar, G. (2024). Threat modeling for cyber warfare against less cyber-dependent adversaries. *Proceedings of the 23rd European Conference on Cyber Warfare and Security*, 21(1), 257-264. doi: 10.34190/eccws.23.1.2462.
[20] Legárd, I. (2020). Building an effective information security awareness program. *Central and Eastern European eDem and eGov Days*, 338, 189-200. doi: 10.24989/ocg.338.15.
[21] Marchenko, O. (2023). Cybersecurity and information protection: Analysis of the impact of risks and threats using modern effective cyberspace protection strategies. *Information Technology: Computer Science, Software Engineering and Cyber Security*, 3, 50-59. doi: 10.32782/IT/2023-3-6.
[22] Matthews, T. (2022). *Creeper: The world's first computer virus*. Retrieved from https://www.exabeam.com/blog/infosec-trends/creeper-the-worlds-first-computer-virus/.
[23] Nemchick, E. (2023). *What is Pegasus spyware+how to remove it from your mobile device?* Retrieved from https://us.norton.com/blog/emerging-threats/pegasus-spyware.
[24] Obi, O.C., Akagha, O.V., Dawodu, S.O., Anyanwu, A.C., Onwusinkwue, S., & Ahmad, I.A.I. (2024). Comprehensive review on cybersecurity: Modern threats and advanced defense strategies. *Computer Science & IT Research Journal*, 5(2), 293-310. doi: 10.51594/csitrj.v5i2.758.
[25] Okhanashvili, S. (2023). Cyber security and malware. *"Intercultural Dialogues" Transactions*, 7. doi: 10.52340/idw.2023.72.
[26] Olaniyan, T. (2021). *Applying the diamond model of intrusion analysis: FireEye breach*. Retrieved from https://www.researchgate.net/publication/354254596_Applying_the_Diamond_Model_of_Intrusion_Analysis_FireEye_Breach.
[27] Qumer, M., & Ikrama, S. (2022). Poppy Gustafsson: Redefining cybersecurity through AI. *The Case for Women*, 1-38. doi: 10.1108/CFW.2022.000001.
[28] Ravi, P., Bhargav, K.S., Venkatesh, P.M.M., Princy, M., & Reddy, M.R. (2024). Empowering security with machine learning for ransomware and malware detection. *International Journal of Scientific Research in Engineering and Management*, 8(4), 1-5. doi: 10.55041/IJSREM30675.
[29] Stallings, W., & Brown, L. (2018). *Computer security: Principles and practice*. London: Pearson.
[30] Tsymbal, A. (2024). *Development of a system for detecting suspicious activities in computer networks*. (Bachelor's thesis, Black Sea National University named after Petro Mohyla, Mykolaiv, Ukraine).
[31] Zulkovska, I., Pluzhnik, A., & Zhulkovski, O. (2021). Modern methods of detection of malware. *Mathematical Modelling*, 44(1), 46-54. doi: 10.31319/2519-8106.1(44)2021.235922.

Розробка шкідливих програм: від ранніх вірусів до сучасних кіберзагроз

Денис Ковальчук
Аспірант
Міжнародний гуманітарний університет
65009, дор. Фонтанська, 33, м. Одеса, Україна
https://orcid.org/0009-0003-2302-8698

Анотація. Шкідливі програми є однією з найбільших загроз у цифровому середовищі, оскільки вони постійно еволюціонують і стають все більш небезпечними. Метою цього дослідження був аналіз еволюції зловмисного програмного забезпечення. Для досягнення цієї мети було застосовано історичний, порівняльний та емпіричний аналіз, а також оцінка існуючих захисних технологій. Основні результати дослідження виявили кілька ключових етапів розвитку шкідливих програм. Історичний аналіз показав, що еволюція шкідливих програм пройшла через кілька значних стадій, від простих вірусів і черв'яків до складних загроз, таких як програми-вимагачі і шпигунські програми. Ці зміни були зумовлені технологічними досягненнями і розширенням можливостей для атак, що дозволило шкідливому програмному забезпеченню використовувати нові вектори впливу та методи обману. Порівняльний аналіз сучасних кіберзагроз розкрив ключові характеристики та відмінності між різними типами шкідливих програм, а також їхні специфічні методи розповсюдження і вразливості. Виявлено, що нові загрози мають складнішу архітектуру і використовують більш інноваційні тактики, що значно ускладнює їхнє виявлення та нейтралізацію. Емпіричний аналіз, що включав використання інструментів виявлення загроз, надав конкретні дані про поведінку шкідливих програм у реальних умовах. Огляд і тестування сучасних методів захисту, включаючи антивірусні рішення, системи виявлення вторгнень та міжмережеві екрани, показали їхні сильні та слабкі сторони, а також їх ефективність у виявленні і запобіганні новим загрозам. Результати дослідження підкреслили необхідність постійного вдосконалення методів захисту, що є критично важливим для ефективної боротьби з сучасними кіберзагрозами

Ключові слова: еволюція інформаційних атак; цифрова безпека; системи виявлення вторгнень; захисні технології; аналіз комп'ютерних інцидентів

BULLETIN of Cherkasy State Technological University

Journal homepage: https://bulletin-chstu.com.ua/en Vol. 29 No. 3. 2024

UDC 004.932.4
DOI: 10.62660/bcstu/3.2024.21

Classification of military equipment based on computer vision methods

Oleh Basystiuk*
Assistant
Lviv Polytechnic National University
79000, 12 Stepan Bandera Str., Lviv, Ukraine
https://orcid.org/0000-0003-0064-6584
Zoriana Rybchak
PhD in Technical Sciences, Associate Professor
Lviv Polytechnic National University
79000, 12 Stepan Bandera Str., Lviv, Ukraine
https://orcid.org/0000-0002-5986-4618
Dmytro Betsa
Student
Lviv Polytechnic National University
79000, 12 Stepan Bandera Str., Lviv, Ukraine
https://orcid.org/0009-0006-6582-8900

Abstract. Means of high-altitude reconnaissance, in particular satellites, reconnaissance drones and aviation complexes, are the most common means for solving the tasks of search and detection of targets. This work focuses on improving the process of finding and identifying targets by implementing an automatic search system using artificial intelligence, with a special emphasis on the use of this technology in drones, under conditions of limited computing resources. The purpose of the work was to create a machine learning model that would localise and classify military equipment using images obtained from unmanned aerial vehicles. Machine learning models used to localise objects in images based on CNN, ResNet, Fast CNN, EfficientDet and YOLO approaches are the research methods. Various computer vision approaches, based on convolutional networks, to localise and classify military equipment in images obtained from unmanned aerial vehicles have been investigated. The approach based on the YOLO8 method has proved to be the most effective one. The generalised precision of the proposed model of image segmentation technique is 70%, and the classification precision is close to 90%, the inference time of the proposed model is less than 400 milliseconds. The system takes an image as input and returns the input image with the found military equipment. In addition, the YOLO8 (nano, small, medium) methods have been tested in the problem of equipment identification and classification in images from unmanned aerial vehicles. The approach proves to be effective and has the potential for further application as well as improvement with larger sets. The system can be used in practice to optimise the search for targets, thus simplifying the task for the operator of unmanned aerial vehicles. Also, in the case of further refinement and optimisation for specific hardware resources, it has the potential for implementation in the real defence sector. Potentially, this solution can become an important tool for military intelligence and other related industries, where precise identification of objects in real-time images is important. The implementation of such systems can significantly increase the efficiency and speed of response in various scenarios of the use of unmanned aerial vehicles

Keywords: localisation; identification; labelling; convolutional neural network; image analysis

Article's History: Received 03.07.2024; Revised 16.09.2024; Accepted 21.10.2024

Suggested Citation:
Basystiuk, O., Rybchak, Z., & Betsa, D. (2024). Classification of military equipment based on computer vision methods. *Bulletin of Cherkasy State Technological University*, 29(3), 21-30. doi: 10.62660/bcstu/3.2024.21.

*Corresponding author

INTRODUCTION

The problem of pattern identification in images is relevant in many industries, including security, medicine, automotive industry and military technology. The identification of military equipment, in particular, has not been sufficiently investigated and described, because it was not of such public and scientific interest, besides, it was a completely non-trivial task. But since the 2000s, humanity has witnessed the rapid technological evolution of object detection and its profound impact on the entire field of computer vision, and therefore object detection, as one of the most fundamental and complex problems of computer vision, is given great attention of scientists.

The article by Z. Zou *et al*. (2023) examines this fast-growing field of research in detail in light of technical evolution spanning more than a quarter of a century (from the 1990s to 2022). The study covers a number of topics, including milestones of detector historical development, datasets, metrics, fundamental building blocks of an identification system, acceleration techniques, and state-of-the-art identification techniques in use. The use of camouflage and active countermeasures during reconnaissance and obtaining results about the equipment are factors that further complicate detection. The ability to precisely identify and classify equipment can be critical to tactical and strategic decisions.

The current state of development of computer vision technologies demonstrates significant achievements due to the rapid development of machine learning and artificial intelligence technologies. J. Liang (2020) has noted that the introduction of residual blocks into the ResNet network allows the network to be further deepened without suffering from performance degradation, i.e. avoiding overtraining. This approach simplifies the training of very deep networks, making possible to build models with hundreds or even thousands of layers, which was previously impractical due to the problem of vanishing gradient. This breakthrough has become a fundamental technique in many complex image identification tasks and has influenced the development of many subsequent deep learning architectures.

It also makes significant progress in solving the problem of pattern identification in images. Recent research focuses on the development of new algorithms and models that can provide high precision and efficiency of object identification in various conditions. In particular, A. Bochkovskiy *et al*. (2020) have focused on the study of activation functions, namely batch normalisation, Weighted-Residual-Connections (WRC), Cross-Stage-Partial-connections (CSP), Cross mini-Batch Normalisation, CmBN), Self-Adversarial-Training (SAT), and have achieved results with an average precision (AP) of 43.5% and an AP50 of 65.7% on the MS COCO dataset, while maintaining real-time processing speed of about 65 frames per second on the Tesla V100.

The article by Z.-Q. Zhao *et al*. (2019) considers an approach to object detection based on deep learning and describes in detail their network architecture, learning strategies, and optimisation functions, in particular, the application of You Only Look Once (YOLO) approaches, covering typical common architectures of object detection, modifications and enhancements to improve performance. In addition, the article considers specific tasks of object detection in the image, in particular, using video. The experimental analysis compares different methods, drawing meaningful conclusions and proposing promising directions and tasks for future research in the field of object detection and neural network-based learning systems.

The work by D. Zhang *et al*. (2024) presents a new imaging paradigm for detecting camouflaged objects using RGB-D images. To accomplish this task, the authors have created a CODD dataset derived from existing RGB-D visible object detection datasets using image-to-image transformation techniques (Yang *et al*., 2023). CODD matches the diversity and complexity of widely used datasets for camouflaged object detection. Researchers G.-P. Ji *et al*. (2023) propose a selection strategy to ensure high image quality, taking into account the structural similarity between images before and after transformation, based on gradient-induced transition, which is a soft grouping between context features and similarity of object appearance and the background, and the fuzziness of object boundaries is also considered.

Identification of military equipment has not been sufficiently investigated in previous research. All of the above studies are oriented towards performing tasks in the civilian sector, and do not include the ability to identify and classify military equipment in the image. Therefore, the purpose of this study was to create a deep learning model for identifying military equipment, taking into account camouflage and counteraction technologies obtained from unmanned aerial vehicles.

MATERIALS AND METHODS

Modern technologies of machine learning and computer vision open up new opportunities for effective monitoring and classification of objects in the image. This article is devoted to a comparative analysis of various deep learning models, including ResNet, CNN, EfficientDet, YOLOv8 (nano, small, medium). The efficiency and possibilities of integration into automated monitoring systems have been investigated.

To work with machine learning methods, the primary task is to collect and label the appropriate dataset. In the case of equipment identification, this is the main means of imaging, since the modern Ukrainian-Russian war, which has been ongoing with varying intensity since 2014, is the largest source of the required data. So, to search for images, it is worth focusing on various social networks, namely Oryx, GeoConfirm and Twitter, where military analysts, journalists, fighters, as well as the brigade show the results of their units' work. All

collected data have been summarised and published in the open access on the platform of Figshare (2024). A generalised order of image processing consists of several main stages, each of which plays an important role in ensuring precise and efficient image processing. The study considers four key steps (Fig. 1).

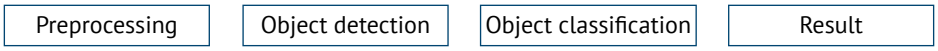

Figure 1. Generalised order of the image processing pipeline by computer vision systems of the proposed system

Source: authors' development

Preprocessing is the initial stage of image processing, which includes image preparation for further analysis. Such image preparation helps to improve its quality and make it more suitable for further processing stages. The main methods used for image preprocessing are: image alignment and normalisation, noise filtering, lighting correction. Object detection is the process of determining the location of objects in the image, that is, in other words, it is the segmentation of the image and the localisation of objects in it. Effective object detection is critical for successful classification and further analysis. Object classification includes identification and determination of the type of object that has been detected in the image at previous steps of the processing pipeline. The classification allows for identification of objects and specific labelling of each of detected objects in the image, according to the classes provided for the search using the designed system. The classification adds a more situational understanding and will also make it possible to better prioritise the tasks of defeating various enemy targets (Benmhahe & Chentoufi, 2021). Results are a process that is responsible for the preparation and correct presentation of the obtained results, which enables the data to be used later by users or other systems for further processing or decision-making.

It is also necessary to determine what types of videos/photos to choose for analysis. In general, first of all, the choice is between night and day videos, as well as between their different types. Optical cameras (standard cameras) are the most common ones and capture optical image (Panwar *et al.*, 2023). Infrared cameras (thermal cameras) capture thermal radiation of objects and build a heat map of the battlefield (Ippalapally *et al.*, 2020). There are many types of images, but they are not practically used in UAVs (unmanned aerial vehicles) and are irrelevant in this study. Infrared cameras capture images in a completely different way than optical cameras (Park & Lee, 2021), in them it is much easier to identify equipment and weapons, so for these cameras the research itself is not very relevant. Images from different types of cameras are shown in Figure 2.

Figure 2. An example of an image from different types of cameras

Source: authors' development

The described leaves only optical cameras. These cameras are mounted on almost all UAVs (Chen *et al.*, 2024), and are the main eyes of the front. And it is for them that this study is the most relevant, since it is quite difficult for the human eye to see the camouflaged equipment, it requires great skills and knowledge. For image processing, this study has used the YOLO model, as it provides a high speed of object real-time detection and classification. Due to its architecture, YOLO processes the image only once, which significantly reduces the computation time.

RESULTS

There are many methods for image processing and object detection in the model used in the study. Convolutional Neural Networks (CNNs) are the basis for many computer vision tasks. They work by training a network on a large number of images, automatically finding and highlighting important features for object classification and detection. Histograms of Oriented Gradients (HOG) are a method for detection of image features to identify objects, which is used to describe the textures and shapes of objects in an image (Zhou *et al.*, 2020). It analyses pixel brightness gradients and creates histograms that indicate the direction and intensity of changes in the image. You Only Look Once (YOLO) is a modern method for simultaneous detection and classification of objects in the image (Hussain, 2023). Speed is the main advantage of YOLO, as it processes the image only once, unlike other

methods that may require multiple passes. For a better understanding, a comparative characteristics and identification precision of various researched methods are given (Table 1).

Table 1. Comparative table of results

Method name	Precision (%)	F1 measure	Model size (MB)
ResNet	66.4	0.86	46
CNN	56.8	0.64	50
EfficientDet	63.4	0.82	32
YOLOv8m	71.8	0.91	54

Notes: the results obtained during the study may differ in other systems and vehicles
Source: authors' development

Additionally, a system deployment diagram that shows how various components interact in a real environment, including hardware and software modules is provided. The proposed system consists of the following components: communication networks, Backend and Database (Fig. 3).

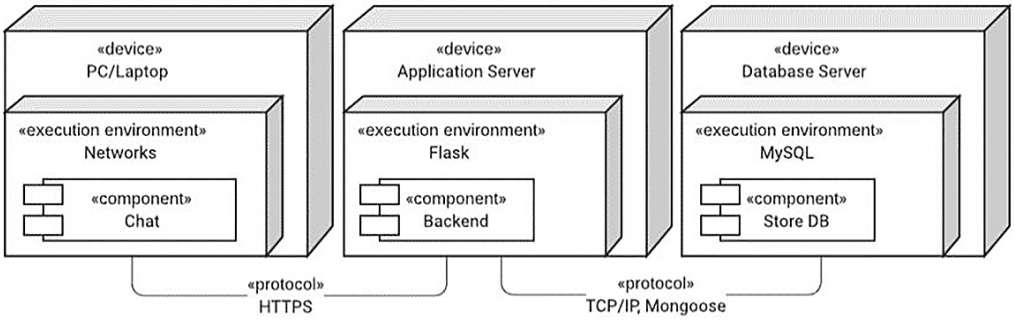

Figure 3. Diagram of the proposed system deployment
Source: authors' development

During the experiments, it has been found that adding of some context to the classification network significantly increases the classification precision and almost eliminates the localisation problem. Thus, the main attention has been focused on the precision of the detection of artillery and military vehicles (Fig. 4).

Class	Images	Instances	Box(P	R	mAP50
all	203	203	0.718	0.61	0.637
artillery	203	30	0.706	0.4	0.43
vehicle	203	173	0.73	0.821	0.845

Figure 4. Precision of the model for two classes of equipment
Source: authors' development

As shown, the precision of artillery detection is significantly lower (more than twice) compared to the precision of military vehicle detection. This can be explained by the quality of images of towed artillery: most images of artillery have low resolution and are fuzzy due to the specificity of its use, namely from closed firing positions and nests. In Figure 5, gun muzzle can be identified on the right, while only a silhouette of the gun is visible on the left, making it difficult to identify the artillery. There is also a problem of sufficient saturation of the dataset with different types of images, in particular with the use of various fortifications and caponiers for equipment. As for military equipment, the results are much better. When testing the neural network in typical images, high results have been obtained. Stroke precision is slightly lower, but this is not critical, since detection precision is the main parameter (Fig. 6).

a

b

Figure 5. Examples of fortification nests of towed artillery
Source: authors' development

Figure 6. Examples of equipment localization in a test dataset
Source: authors' development

The obtained model shows high results of identification precision, but there is a problem with sufficient amount of data. There is a big problem with the variety of equipment forms in the images and with the variety of environments. At first glance, the network appears to find most equipment, but in fact it only detects it in certain image formats. Manual testing has shown that the network works well with images where equipment is not too dark and where certain structural elements are visible. This is due to insufficient number of similar images in the dataset. The problem has been partially solved by increasing the number of such images in the dataset. Currently, the model finds about 60% of small objects. Other similar problems have been corrected by analysing the images and supplementing the dataset. For example, an image of tanks from the side projection has been added (Fig. 7).

Figure 7. Results of equipment classification from the side projection
Source: authors' development

During the study, it has been found that smoke is an important artifact that significantly impairs the performance of the identification model. Not only does smoke reduce the precision of equipment identification in images where it is present, but often thick clouds of smoke are classified as equipment. This creates significant obstacles for precise identification of military objects, which is especially critical in the conditions of hostilities. Examples of identification of a military equipment convoy in the video of the movement of a convoy of vehicles in a smoke screen demonstrate these difficulties.

There are differences between civilian and military vehicles that affect the process of their identification. Military equipment is often camouflaged in every possible way to minimise its detection by reconnaissance devices, especially drones. This is done with the help of various camouflage nets, forest plantations or by hiding equipment in dense terrain. Civilian vehicles are characterised by distinctive features such as visible wheels and bodywork, which make them easy to identify, even when they are partially covered. In the case of military equipment, there are no clear patterns for identification; often only individual pieces of equipment are visible, such as a barrel that can belong to either a tank gun or towed artillery, or a turret that can belong to either an infantry fighting vehicle (IFV) or a tank. This greatly complicates the identification of the type of equipment.

Additional difficulties are created by active countermeasures from the enemy, who tries to damage the operation of reconnaissance vehicles and shoot them down, in particular, UAVs. For this, means of Electronic Warfare (EW) and Air Defence (AD) are used, which significantly limits the efficiency of UAVs. EW tools are used to create electronic interference that can affect the operation of navigation systems, communication systems and UAV control (Zhang *et al.*, 2024). This can result in loss of communication with the vehicle, its disorientation or even complete loss of control over it.

In view of this, the study has been aimed at developing two models: classification and localisation ones, which have been considered separately. The localisation model has two classes to search for: vehicle and artillery. The YOLO8s neural model, which has demonstrated optimal identification results, has been chosen. The results of this model are presented in detail in Figure 8.

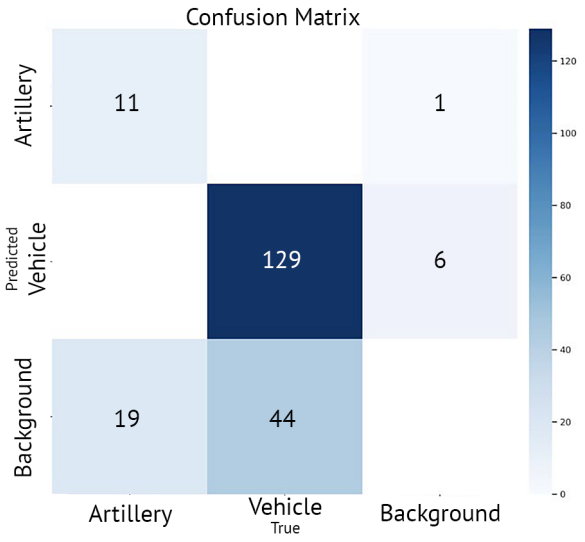

Figure 8. Matrix of uncertainties of the proposed YOLO model
Source: authors' development

The obtained results show that the proposed model has achieved a precision of 71.8%, which exceeds

the indicators of other considered models. The value of F1-measure for the proposed model has been 0.91, which indicates a high balance between the precision and completeness of identification (Fig. 9).

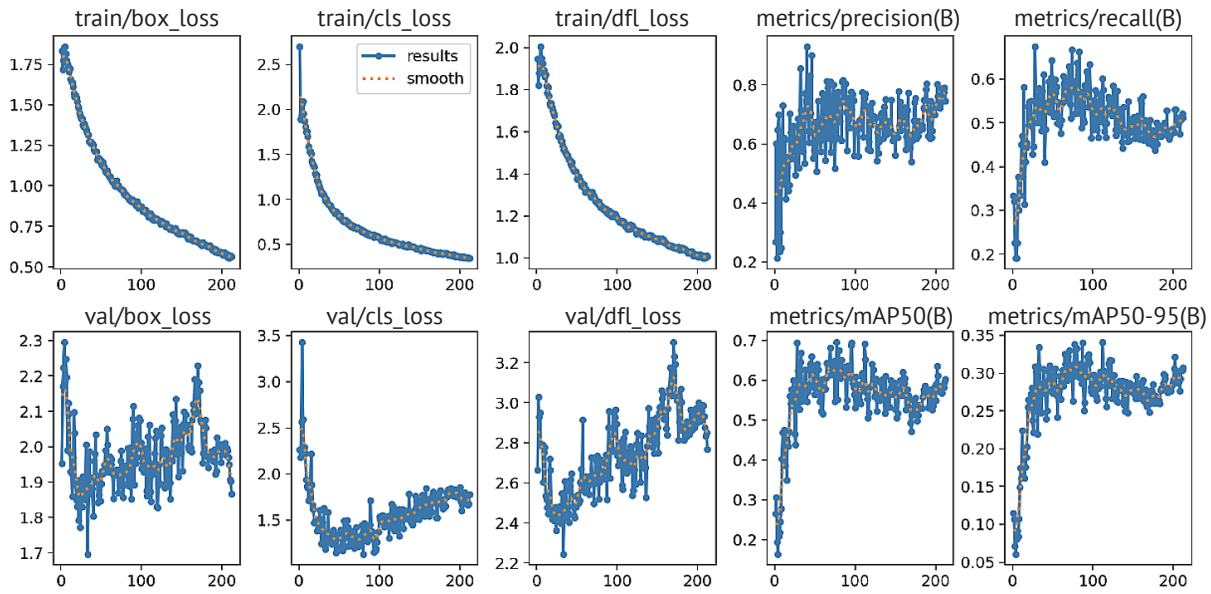

Figure 9. Results of the proposed YOLO model training

Source: authors' development

Precision is based on two indicators: precision of vehicle or artillery identification, as well as precision of equipment boundaries. For this work, the precision of classification is much more important than that of localisation. The results of the study confirm that the model is able to identify fragments of military equipment, such as artillery barrels and turrets of combat vehicles, which is important for precise identification of the type of equipment.

The proposed model demonstrates a high level of precision and efficiency in the tasks of identifying military equipment under camouflage conditions. The results of the study show that the model can be used to improve the effectiveness of intelligence operations and provide more precise detection of threats, which is critical for ensuring national security and defence.

DISCUSSION

The model proposed in this study will combine modern deep learning methods with new algorithms to increase the precision and reliability of military objects detection and classification. This is especially true in conditions where the enemy uses various methods to complicate the detection of its equipment, including EW and AD equipment.

It is worth noting that the images obtained from a multispectral camera capture images in different ranges (visible light, ultraviolet radiation, etc.). A comprehensive study on the topic of multispectral imaging has been conducted by V. Mohammadi *et al.* (2022). The authors consider a camera sensor consisting of a multispectral image sensor, a spectral filter matrix, a sensor board, and a control board. A hybrid system that operates using eight bands in the visible range, in the interval of 400-700 nm is proposed. Based on a genetic algorithm, a program has been developed to find the best combination of filters. The program selects Gaussian filters using the genetic algorithm based on the Wiener filter estimation method. For band selection, a minimum RMS of 0.0016 has been obtained to select bands in the visible range. The developed camera provides eight high-resolution spectral images.

The article by L. Alzubaidi *et al.* (2021) describes the sequence of CNNs development, starting with the AlexNet network and ending with the High-Resolution Network (HR.Net). Also a list of problems and standardised solutions that have helped researchers to better understand existing research gaps by applying pre-defined techniques and approaches is provided. In addition, a list of the main areas of application of the proposed approaches is provided, and hardware requirements for each of the proposed models are given. Considering the further development of convolutional neural networks, namely the EfficientDet approach, M. Tan *et al.* (2020) propose a weighted bidirectional feature pyramid network (BiFPN) that allows easy and fast fusion of multi-scale features, and a combined scaling method that uniformly scales resolution, depth, and width for all base networks, feature networks, and class prediction networks simultaneously. Collectively, all these improvements have resulted in the development of a new family of object detectors, called EfficientDet, which consistently achieve much better efficiency. In particular, when using the EfficientDet-D7

model to compare with convolutional networks, the former achieves 52.2 AP on the COCO test device with 52M and 325B FLOPs1, being 4-9 times smaller and using 13-42 times less FLOPs than that of CNN (Kishore & Balamurugan, 2024).

G. Csurka *et al.* (2022) summarise two decades of image segmentation research. Since current segmentation models require large amounts of annotated data, which are expensive, the authors highlight the success of Unsupervised Domain Adaptation (UDA), which allows for a significant reduction in data labelling costs by using domain-based image clustering. A review of five years of experience with Domain Adaptation for Semantic Image Segmentation (DASiS), which highlights the importance of adapting segmentation models to new conditions, is the second key contribution. Finally, commonly used datasets and benchmarks are described, and supporting machine learning techniques, namely the latest trend in the use of transformers, are discussed. Another method involves the use of the Gray Level Co-occurrence Matrix (GLCM) method, which has been analysed in detail in the work of N. Iqbal *et al.* (2021). The GLCM method, analysing spatial relationships between pixels, allows for extraction of valuable information for such tasks as texture classification and segmentation. Experimenting with various parameters and functions, it is possible to fine-tune the analysis to meet specific application requirements, which makes this method a universal tool in the field of image processing and computer vision.

R. Gavrilescu *et al.* (2018) consider the Faster Regional based Convolutional Neural Network (Faster R-CNN) algorithm. Faster R-CNN is the result of merging of Region Proposal Network (RPN) and Fast-RCNN algorithms into one network. To increase video processing power, a Graphics Processing Unit (GPU) has been used to train and test real-time object identification at 15 frames per second in a dataset containing 3000 images for 4 classes. The dataset consists of images containing three traffic light phases. The use of the You Only Look Once (YOLO) approach, which has been also used in the above study, and its advantages and disadvantages have been described by M. Sohan *et al.* (2024). The authors have reviewed the YOLOv8 model, which is an improved version for real-time object detection. Of all popular methods for object identification and machine learning models, such as Faster RCNN, SSD, and RetinaNet, YOLO is the most well-known method in terms of precision, speed, and efficiency. This study presents an analysis of YOLO v8, highlighting its innovative features, improvements, applicability in various environments, and a detailed comparison of its performance indicators with other versions and models. Additionally, the authors A. Vijayakumar & S. Vairavasundaram (2024) describe the stages of development of deep learning methods over the decades, emphasise the relevance of the topic, and state that most researchers are working on improving the detection, segmentation and classification of objects.

In the conducted research, the application of technology for military equipment identification, which has not been proposed in any of the reviewed studies, is a unique feature of the final development. This technology is based on the integration of modern approaches to image analysis, which allows for effective identification of military equipment in conditions of limited computing resources. The approach uses advanced deep learning techniques, including neural networks optimised for use in unmanned aerial vehicles. During the research, an effective selection strategy to ensure high image quality has been found. This strategy takes into account the structural similarity between images before and after transformation, based on gradient-induced transition. This approach provides a soft grouping between context features and similarity of object appearance and the background. Taking into account of the fuzziness of object boundaries within the framework of the developed system allows for increase in the precision of identification even in difficult conditions.

The use of an image paradigm to detect camouflaged objects using RGB-D images is one of key innovations in the conducted research. This approach allows for a significant improvement in the precision of camouflaged equipment detection, as it includes additional information about the depth of the scene. The use of RGB-D images in combination with developed data processing algorithms allows for effective separation of camouflaged objects from the background, which is important in the context of military applications. The proposed system has been tested in an optical camera with 24 frames per second, with the possibility of increasing to 60, and a response time within 400 milliseconds. These characteristics provide a significant advantage compared to other considered systems, especially in complex conditions of battlefield images and when detecting camouflaged equipment. The high frame rate and fast response time allow the system to work in real time, which is critical for prompt detection and identification of military objects.

CONCLUSIONS

Within this study, a wide range of architectural solutions based on neural networks for searching for objects in images has been analysed. In particular, various approaches to solving the task of finding equipment in the image have been considered, ultimately choosing localisation and identification. Various approaches for object identification, including one-stage and two-stage methods, have been investigated. A dataset from identified sources has been collected and models for object localisation and classification using YOLOv8 have been trained.

Despite significant limitations in the amount of data, the models show improved results compared to known and publicly available solutions. For image formats well represented in the dataset, the network

demonstrates 80% of localisation precision and 70% of classification precision. Overall, this approach has proven to be effective and has the potential to be improved with more data that will make it possible to more precisely locate, segment, and classify the equipment in images. This approach to solving the problem has proven itself well, even despite significant limitations in the dataset. High precision of identification and classification of most types of images has been achieved.

The main conclusion is that if there is a sufficiently large dataset, it is possible to create a neural network that will effectively and reliably identify equipment at the battlefield. Further research and improvement of the machine learning model will take place precisely in the direction of supplementing and retraining the model on larger (newer) datasets, as well as in the way of integrating additional sources of information with various, in particular thermal, multispectral images, as well as various, in particular audio, tracks with the sound of equipment and battlefields.

ACKNOWLEDGEMENTS

None.

CONFLICT OF INTEREST

None.

REFERENCES

[1] Alzubaidi, L., *et al*. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8, article number 53. doi: 10.1186/s40537-021-00444-8.

[2] Benmhahe, B., & Chentoufi, J.A. (2021). Automated pavement distress detection, classification and measurement: A review. *International Journal of Advanced Computer Science and Applications*, 12(8), 708-718.

[3] Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y.M. (2020). YOLOv4: Optimal speed and accuracy of object detection. *arXiv*, 2004, article number 10934. doi: 10.48550/arXiv.2004.10934.

[4] Chen, X., Liu, C., Chen, L., Zhu, X., Zhang, Y., & Wang, C. (2024). A pavement crack detection and evaluation framework for a UAV inspection system based on deep learning. *Applied Sciences*, 14(3), article number 1157. doi: 10.3390/app14031157.

[5] Csurka, G., Volpi, R., & Chidlovskii, B. (2022). Semantic image segmentation: Two decades of research. *Foundations and Trends® in Computer Graphics and Vision*, 14(1-2), 1-162. doi: 10.1561/0600000095.

[6] Figshare. (2024). *Vehicles and fortification image recognition dataset*. doi: 10.6084/m9.figshare.25726470.

[7] Gavrilescu, R., Zet, C., Foșalău, C., Skoczylas, M., & Cotovanu, D. (2018). Faster R-CNN: An approach to real-time object detection. In *International conference and exposition on electrical and power engineering (EPE)* (pp. 165-168). Iasi: IEEE. doi: 10.1109/ICEPE.2018.8559776.

[8] Hussain, M. (2023). YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection. *Machines*, 11(7), article number 677. doi: 10.3390/machines11070677.

[9] Ippalapally, R., Mudumba, S.H., Adkay, M., & Vardhan, H.R.N. (2020). Object detection using thermal imaging. In *2020 IEEE 17th India Council international conference (INDICON)* (pp. 1-6). New Delhi: IEEE India Council. doi: 10.1109/INDICON49873.2020.9342179.

[10] Iqbal, N., Mumtaz, R., Shafi, U., & Zaidi, S.M.H. (2021). Gray level co-occurrence matrix (GLCM) texture based crop classification using low altitude remote sensing platforms. *PeerJ Computer Science*, 7, article number e536. doi: 10.7717/peerj-cs.536.

[11] Ji, G.-P., Fan, D.-P., Chou, Y.-C., Dai, D., Liniger, A., & Gool, L.C. (2023). Deep gradient learning for efficient camouflaged object detection. *Machine Intelligence Research*, 20, 92-108. doi: 10.1007/s11633-022-1365-9.

[12] Kishore, A.S.K., & Balamurugan, G. (2024). Faster region based convolution neural network with context iterative refinement for object detection. *Measurement: Sensors*, 31, article number 101025. doi: 10.1016/j.measen.2024.101025.

[13] Liang, J. (2020). Image classification based on RESNET. *Journal of Physics: Conference Series*, 1634, article number 012110. doi: 10.1088/1742-6596/1634/1/012110.

[14] Mohammadi, V., Sodjinou, S.G., Katakpe, K.K., Rosse, M., & Gouton, P. (2022). Development of a multi-spectral camera for computer vision applications. In *30. jubilee international conference in central Europe on computer graphics, visualization and computer vision* (pp. 24-27). Plzen: Vaclav Skala University of West Bohemia. doi: 10.24132/CSRN.3201.4.

[15] Panwar, K., Singh, A., Kukreja, S., Singh, K.K., Shakhovska, N., & Boichuk, A. (2023). Encipher GAN: An end-to-end color image encryption system using a deep generative model. *Systems*, 11(1), article number 36. doi: 10.3390/systems11010036.

[16] Park, S., & Lee, H. (2021). Deep learning approach to optical camera communication receiver design. In *2021 IEEE region 10 symposium (TENSYMP)* (pp. 1-5). Jeju: IEEE. doi: 10.1109/TENSYMP52854.2021.9550896.

[17] Sohan, M., Sai Ram, T., & Rami Reddy, C.V. (2024). A review on YOLOv8 and its advancements. In I.J. Jacob, S. Piramuthu & P. Falkowski-Gilski (Eds.), *Data intelligence and cognitive informatics* (pp. 529-545). Singapore: Springer. doi: 10.1007/978-981-99-7962-2_39.

[18] Tan, M., Pang, R., & Le, Q.V. (2020). EfficientDet: Scalable and efficient object detection. In *IEEE/CVF conference on computer vision and pattern recognition (CVPR)* (pp. 10778-10787). Seattle: IEEE. doi: 10.1109/CVPR42600.2020.01079.
[19] Vijayakumar, A., & Vairavasundaram, S. (2024). YOLO-based object detection models: A review and its applications. *Multimedia Tools and Applications*, 83, 83535-83574. doi: 10.1007/s11042-024-18872-y.
[20] Yang, S., Jiang, L., Liu, Z., & Loy, C.C. (2023). GP-UNIT: Generative prior for versatile unsupervised image-to-image translation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(10), 11869-11883. doi: 10.1109/TPAMI.2023.3284003.
[21] Zhang, D., Wang, C., & Fu, Q. (2024). A new benchmark for camouflaged object detection: RGB-D camouflaged object detection dataset. *Open Physics*, 22(1), article number 20240060. doi: 10.1515/phys-2024-0060.
[22] Zhang, Z., Xie, X., Guo, Q., & Xu, J. (2024). Improved YOLOv7-Tiny for object detection based on UAV aerial images. *Electronics*, 13(15), article number 2969. doi: 10.3390/electronics13152969.
[23] Zhao, Z.-Q., Zheng, P., Xu, S.-T., & Wu, X. (2019). Object detection with deep learning: A review. *IEEE Transactions on Neural Networks and Learning Systems*, 30(11), 3212-3232. doi: 10.1109/TNNLS.2018.2876865.
[24] Zhou, W., Gao, S., Zhang, L., & Lou, X. (2020). Histogram of oriented gradients feature extraction from raw bayer pattern images. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 67(5), 946-950. doi: 10.1109/TCSII.2020.2980557.
[25] Zou, Z., Chen, K., Shi, Z., Guo, Y., & Ye, J. (2023). Object detection in 20 years: A survey. *Proceedings of the IEEE*, 111(3), 257-276. doi: 10.1109/JPROC.2023.3238524.

Класифікація військової техніки на основі методів комп'ютерного зору

Олег Басистюк
Асистент
Національний університет «Львівська політехніка»
79000, вул. Степана Бандери, 12, м. Львів, Україна
https://orcid.org/0000-0003-0064-6584
Зоряна Рибчак
Кандидат технічних наук, доцент
Національний університет «Львівська політехніка»
79000, вул. Степана Бандери, 12, м. Львів, Україна
https://orcid.org/0000-0002-5986-4618
Дмитро Беца
Студент
Національний університет «Львівська політехніка»
79000, вул. Степана Бандери, 12, м. Львів, Україна
https://orcid.org/0009-0006-6582-8900

Анотація. Найбільш поширеними засобами для вирішення завдань пошуку та виявлення цілей є засоби висотної розвідки, зокрема супутники, розвідувальні дрони та авіаційні комплекси. Дана робота фокусується на покращенні процесу знаходження та ідентифікації цілей шляхом імплементації системи автоматичного пошуку із використанням штучного інтелекту, з особливим акцентом на використання цієї технології на дронах, в умовах обмежених обчислювальних ресурсів. Метою роботи було створення моделі машинного навчання, яка дозволить локалізувати та класифікувати військову техніку за допомогою зображень, отриманих з безпілотних літальних апаратів. Методами дослідження є моделі машинного навчання, що використовуються для локалізації об'єктів на зображеннях, що базуються на підходах CNN, ResNet, Fast CNN, EfficientDet та YOLO. Було досліджено різні підходи комп'ютерного зору, на основі згорткових мереж для локалізації та класифікації військової техніки на зображеннях отриманих з безпілотних літальних апаратів. Найефективніше себе показав підхід, який базується на методі YOLO8. Узагальнена точність пропонованої моделі сегментації техніки на зображення становить 70 %, а точність класифікації наближається до 90 %, inference time пропонованої моделі становить менше 400 мілісекунд. Система приймає як вхідні дані зображення і повертає вхідне зображення із знайденою військовою технікою. Додатково проведено апробацію методів YOLO8 (nano, small, medium) у проблематиці розпізнавання та класифікації техніки на зображеннях з безпілотних літальних апаратів. Підхід виявився ефективним і має потенціал для подальшого застосування, а також покращення при наявності більших наборів. Система може використовуватися на практиці для оптимізації пошуку цілей, спрощуючи таким чином завдання для оператора безпілотних літальних апаратів. Також, у випадку подальшого доопрацювання та оптимізації під конкретні апаратні ресурси, має потенціал до впровадження у реальному секторі оборони. Потенційно дане вирішення може стати важливим інструментом для військової розвідки та інших суміжних галузей, де важлива точна ідентифікація об'єктів на зображеннях в реальному часі. Впровадження таких систем може значно підвищити ефективність і швидкість реагування в різних сценаріях використання безпілотних літальних апаратів

Ключові слова: локалізація; розпізнавання; маркування; згорткова нейронна мережа; аналіз зображень

BULLETIN of Cherkasy State Technological University

Journal homepage: https://bulletin-chstu.com.ua/en Vol. 29 No. 3. 2024

UDC 004.056.55:004.75
DOI: 10.62660/bcstu/3.2024.31

Data encryption as a method of protecting personal data in a cloud environment

Samur Ahmadov*
Postgraduate Student
Azerbaijan Technical University
AZ1073, 25 Huseyn Cavid Ave., Baku, Azerbaijan
https://orcid.org/0000-0003-0733-898X

Abstract. In the context of cloud technologies, encryption plays a critical role, as data are constantly transmitted over the network and stored on remote servers, which makes them a potential target for cyber-attacks. The purpose of the study lied in a comprehensive analysis of data encryption methods as the main tool for protecting personal information in cloud services. Modern encryption technologies, including symmetric and asymmetric encryption, and their application in various cloud platforms were considered. A comparative analysis of these methods was conducted in terms of their effectiveness, impact on system performance, and complexity in implementation. An important aspect of the study was the examination of problems related to the management of encryption keys, including their secure storage and protection from unauthorised access. The study also examined examples of successful encryption implementation on popular cloud platforms and ways to ensure their compliance with the requirements of legislation in the field of personal data protection. The regulatory acts regulating the processing and storage of personal information and their impact on the choice and implementation of encryption methods in the cloud were analysed. The results of the study showed that encryption remains one of the most reliable ways to protect data in a cloud environment but an integrated approach is needed for its effective application. Optimal data protection includes not only encryption but also key management, regular security monitoring, and staff training. This helps minimise the risks of data leaks and increase user confidence in cloud services

Keywords: key management; system performance; information security; information leaks; confidentiality

INTRODUCTION

With the rapid development of information technology and an increase in the volume of data transmitted and stored in cloud services, personal data protection has become one of the most important tasks for organisations at all levels. Many companies face cyber-attacks, and many of them recognise that the leakage of personal data can substantially damage their reputation and financial position. In this regard, data encryption seems to be a critical tool for ensuring the confidentiality and security of personal information.

The relevance of the topic of data encryption in the cloud environment is due not only to an increase in the number of cyber-attacks but also to changes in legislation requiring organisations to strictly comply with personal data protection standards. Failure to comply with these requirements can lead to serious fines and loss of trust on the part of users. In the context of globalisation and an increasing number of international transactions, companies are faced with the need to comply with various legal norms, which complicates the process of data protection.

Article's History: Received 23.06.2024; Revised 05.09.2024; Accepted 21.10.2024

Suggested Citation:
Ahmadov, S. (2024). Data encryption as a method of protecting personal data in a cloud environment. *Bulletin of Cherkasy State Technological University*, 29(3), 31-41. doi: 10.62660/bcstu/3.2024.34.

*Corresponding author

The problem of data encryption lies not only in its implementation but also in the need to ensure a balance between security and system performance. The use of complex encryption algorithms can negatively affect the performance of cloud services, which requires developers to find optimal solutions. In addition, many organisations face a lack of knowledge and resources to properly implement encryption, which leads to vulnerabilities in their information systems.

The basics of data encryption cover a variety of methods and algorithms used to protect information. These methods allow converting data into an unreadable format that can only be decrypted using the appropriate keys. However, despite its advantages, encryption is not a universal solution; it must be integrated into an overall data security strategy that includes authentication, access control, and regular audits.

Data encryption in a cloud environment allows identifying several critical studies that contribute to understanding the issues of personal data protection. H. Abroshan (2021) investigated the impact of encryption on the performance of cloud services and concluded that the use of encryption algorithms can lead to a noticeable increase in request processing time. The author suggested balancing the level of security and performance, which is critical for the successful implementation of encryption. M. Zeng *et al*. (2019) focused on the importance of asymmetric encryption in cloud applications. They showed that this methodology provides a high level of security when exchanging data between users and cloud providers, reducing the risk of information interception. B. Custers *et al*. (2019), in turn, described approaches to integrating data encryption into an access control system. They noted that the use of encryption in combination with reliable authentication substantially increases the level of protection of personal data. V. Rudnytsky *et al*. (2022) also conducted a review of modern encryption algorithms and offered a comparative analysis of their effectiveness and security. They determined that some new algorithms provide a higher level of protection at a lower cost of resources. M.N. Asghar *et al*. (2019) investigated the problems associated with the legislative regulation of data encryption. They argued that many companies do not comply with the requirements of the General Data Protection Regulation (GDPR) (2016) due to the lack of clear guidance on the use of encryption.

H. Malvai *et al*. (2023) drew attention to the need to ensure transparency in encryption processes. They argued that users should be aware of how their data is encrypted and stored to build trust in cloud providers. D.K.R. Shukla *et al*. (2021), in turn, analysed the practical aspects of implementing encryption in cloud services. Their results showed that best practices include regular updating of encryption keys and the use of multi-factor authentication. K.R. Sajay *et al*. (2019) focused on comparing different encryption protocols for data transmission over cloud services. Their paper has established that using Secure Sockets Layer or Transport Layer Security is the most secure solution to protect data during transmission. T. Palit *et al*. (2019) investigated the role of encryption in protecting data from internal threats. They argued that data encryption at the storage level is an important measure against leaks related to the actions of unscrupulous employees. A. Calder & S. Watkins (2024) emphasised the need for a comprehensive approach to data security, including both technical and organisational measures.

An analysis of the papers of these authors shows that data encryption is a multifaceted topic that requires an integrated approach. Despite a substantial amount of research in the field of data encryption in the cloud environment, there are still subjects that are insufficiently examined or require more in-depth study. One of these subjects is the optimisation of encryption algorithms to improve performance while maintaining a high level of security. In addition, the issues of encryption key management in multi-cloud environments and the problems of compliance with international norms and standards when working with personal data on a global scale remain insufficiently investigated. The purpose of the study was to explore the best approaches to using data encryption in a cloud environment, considering the need to improve performance and comply with international security standards. The objectives of the study were to analyse modern encryption algorithms in terms of their performance and security in cloud services and analyse existing encryption key management methods and their applicability in multi-cloud environments.

MATERIALS AND METHODS

In this study, various encryption methods and algorithms were used to protect both personal and corporate information and to comply with GDPR requirements. Countries that have implemented regulations similar to the EU GDPR were selected for the study. The main objects of the analysis were the legal systems of Argentina, Japan, Brazil, Canada, and South Korea, which have implemented data protection laws (Data privacy laws…, 2024).

The materials for the study were the texts of these legislative acts, and documents of the European Commission establishing criteria for assessing the adequacy of data protection. The methodology included comparative legal analysis, classification of requirements, and identification of differences in the regulation of personal data processing and protection. Indicators such as user information requirements, data processing consent management, data subjects' rights and information security mechanisms were used for the analysis. The first stage of this study was the selection of suitable encryption algorithms. Various methods have been examined that have a high degree of safety and are used

in real conditions to do this. During the selection process, special attention was paid to algorithms such as Advanced Encryption Standard (AES) and Rivest-Shamir-Adleman (RSA). AES, which is a symmetric algorithm, has proven itself to be a high-performance and reliable tool widely used in commercial applications. RSA, being an asymmetric method, plays an important role in the secure transfer of keys and digital signatures, which makes it indispensable in data protection conditions.

Three cloud platforms were considered for the study: Amazon Web Services (2024) (AWS), Microsoft Azure (Azure encryption overview, 2024), and Google Cloud Platform (GCP) (Key purposes and..., 2024). Each of them implements both symmetric and asymmetric encryption mechanisms to protect data. AWS used Amazon S3 Server-Side Encryption (SSE) and AWS Key Management Service (KMS), which use AES-256 to encrypt data. In Microsoft Azure, the Azure Storage Service Encryption system and Azure Key Vault for key management were analysed. GCP explored Google Cloud KMS for symmetric encryption and support for asymmetric methods for secure key transfer.

The next important step was to study the encryption keys, which is a critical aspect of data protection. Both static and dynamic keys were investigated. Static keys ensure consistency throughout the lifetime, while dynamic keys are created for each session, which substantially reduces the risk of compromise. Various metrics were examined to analyse the performance of the encryption algorithms used. Primarily, this is the encryption and decryption time, which allows understanding how fast algorithms can process data. In addition, a security assessment was conducted based on known vulnerabilities and attack resistance. Equally important was the examination of the use of resources, such as RAM and processor time, in the process of performing encryption operations.

At the final stage of the study, a comparative analysis of the data was conducted. This allows not only to compare the execution time of various algorithms but also to assess their safety based on theoretical and practical scenarios. As a result of the analysis, it became clear that the choice of encryption methods, algorithms and approaches to key generation and management directly affects the level of protection of personal information. The correct application of these technologies not only contributes to compliance with GDPR requirements but also ensures a high level of security in conditions of a constant increase in the volume of processed data.

RESULTS

Data encryption plays a critical role in modern digital security. The volume of processed and stored data is growing every year, which makes the need to protect personal and corporate information more urgent. Encryption is the process of converting information into an unreadable format to protect against unauthorised access. In addition to its technical significance, encryption also complies with modern regulatory requirements such as GDPR.

The main purpose of encryption is to ensure data confidentiality. This means that only authorised users can access encrypted information. In addition, encryption guarantees the integrity of the data, preventing it from being altered or corrupted during transmission. In modern conditions, encryption also helps companies comply with regulatory requirements, which is especially important for protecting users' personal information. GDPR, for example, is an EU regulation adopted on April 14, 2016, and introduced on May 25, 2018, to protect the personal data of EU citizens, regardless of where they are processed.

Data encryption is an important element of protecting both personal and corporate information, preventing data leaks and theft. The encryption process involves choosing an algorithm that converts the source data into an encrypted format using a key. The encryption algorithm and the key work together to make the data unreadable to unauthorised persons. This process includes several stages, starting with the selection of an algorithm and ending with the generation of an encryption key. Some algorithms provide static keys, while others are dynamically updated in each session (Rafique *et al.*, 2021).

There are two main types of encryptions: symmetric and asymmetric. Symmetric encryption uses a single key for both encryption and decryption of data. This is faster and requires fewer resources compared to asymmetric encryption, but requires secure key transfer between the parties, which is its weak point. Asymmetric encryption, on the other hand, uses a key pair: public and private. The public key is used to encrypt data and the private key is used to decrypt it, which provides a higher level of security, although the process is slower (Marqas *et al.*, 2020). Table 1 shows a comparison of symmetric and asymmetric encryption.

Table 1. Comparison of symmetric and asymmetric encryption

Characteristics	Symmetric encryption	Asymmetric encryption
Using keys	The same key is used for encryption and decryption. Both the sender and the recipient must own this key	The encryption uses a public key that is accessible to everyone. For decryption – a private key known only to the recipient
Key transfer	It is necessary to securely transfer the key from the sender to the recipient to successfully exchange information. The security of key transfer is the weak point of the method	The public key can be opened and transferred without risk. Only the private key should be kept secret, which makes key transfer more secure

Continued Table 1.

Characteristics	Symmetric encryption	Asymmetric encryption
Efficiency	Encryption is fast and requires less computing resources, which makes it more efficient to work with large amounts of data	Encryption is slower and requires more computing resources, especially at the decryption stage. Suitable for smaller amounts of data and more sensitive operations
Usage examples	It is used for encrypting files, databases, data transmission over the network, and other situations where fast processing and encryption of large amounts of data is required. For example, to encrypt file systems or stream data	It is used to protect key exchange, digital signatures, and data in e-mail, online transactions and other cases where the secure transfer of confidential information is required. For example, to securely exchange symmetric encryption keys over the Internet
Key security	The main threat is the interception of the key since its knowledge gives the attacker the opportunity to decrypt all data	The public key is accessible to everyone, and even if it is intercepted, an attacker will not be able to decrypt the data without the private key. The private key is protected and must remain secret
Algorithms	Examples: AES, Data Encryption Standard	Examples: RSA, Elliptic Curve Cryptography
Visual concept	Visually, the symmetric encryption process can be represented as the interaction of two parties (sender and recipient) who use the same key to encrypt and decrypt messages. Both participants must agree on this key in advance, and it must remain secret from outsiders	Visually, asymmetric encryption can be depicted as a process in which the sender encrypts a message using a public key (which is available to everyone), and the recipient decrypts it using a private key that is kept secret. Thus, the security of key transfer is not violated
Combination of methods	It is used in combination with asymmetric encryption to encrypt large amounts of data after secure key exchange through asymmetric methods	It is used for the initial secure exchange of symmetric encryption keys, which allows using faster symmetric algorithms for data transmission in the future

Source: compiled by the author

Symmetric and asymmetric encryption are widely used on various cloud platforms to protect data during storage and transmission. In AWS, symmetric encryption is implemented through Amazon S3 SSE and AWS KMS, which use AES-256 to quickly encrypt large amounts of data, while asymmetric encryption is used for key management and digital signatures. In Microsoft Azure, symmetric encryption is provided using Azure Storage Service Encryption, which automatically encrypts data, and for asymmetric encryption, Azure Key Vault is used to manage SSL/TLS certificates. On GCP, data is automatically encrypted with symmetric algorithms such as AES-256 via Google Cloud KMS and asymmetric encryption is supported for secure key transfer and digital signatures. Thus, symmetric encryption provides high performance for encrypting large amounts of data, while asymmetric encryption is used for specialised tasks such as secure key transfer and digital signatures.

Symmetric encryption is especially popular due to its simplicity and speed. It is often used in systems where fast data processing is required, such as databases and cloud systems. Algorithms like AES provide reliable encryption that requires minimal computing resources, which makes them attractive for large amounts of data. However, symmetric encryption has its drawbacks, including the need to securely transfer the secret key, which can be a problem when working with multiple users (Gui *et al.*, 2023). In large organisations, key management becomes a complex task that requires regular updating and key protection.

Symmetric encryption, although it is an effective and fast method of data protection, is not without drawbacks. One of the main problems of this approach is the need to securely transfer the secret key between the participants. If the key falls into the hands of intruders, they will be able to decrypt all encrypted data. This makes the key exchange process vulnerable and requires the use of additional security measures, such as secure data channels. Also, key management in large organisations can be difficult, especially when it requires regular key updates for multiple users, which increases the risk of leakage (Li *et al.*, 2019).

Asymmetric encryption is a more complex but secure encryption method. Its key feature is that the public key can be freely distributed, and the private key is kept secret. This solves the problem of secure key transfer since the public key can be accessed by any user who wishes to send an encrypted message without worrying about its transmission through secure channels. However, asymmetric encryption requires more computing resources, which makes it less efficient for large amounts of data. The process of asymmetric encryption begins with the generation of a key pair. The public key can be freely distributed, and the private key is kept by the owner (Senthilkumar & Geetha, 2020). When the sender wants to encrypt a message, they use the recipient's public key. After encryption, the data is

converted into ciphertext and transmitted over any connection, even if it is insecure. Even if an attacker intercept encrypted data, they will not be able to decrypt it without a private key.

The key advantage of asymmetric encryption is its ability to securely transfer keys over open networks. This makes it especially important for applications such as e-commerce, banking transactions, and systems where data must remain confidential. However, the slowness of this method makes it impractical for large amounts of data. One of the main challenges when using encryption is key management. Organisations need to develop reliable systems for creating, storing, transferring, and updating encryption keys. Loss or compromise of the key may result in loss of access to the data or its disclosure. This is especially important for companies that work with confidential information, such as financial data or personal information of users.

It is advisable to consider in more detail the disadvantages of asymmetric encryption. Firstly, it is usually slower than symmetric encryption due to the more complex mathematical operations used in the encryption and decryption process. Additionally, performing asymmetric encryption requires more computing resources, which can be a problem for devices with limited capabilities. Managing key pairs can also be difficult, especially in large organisations where many users interact with each other. Nevertheless, asymmetric encryption is a powerful tool for ensuring data security in modern digital communications. Due to the use of a key pair, it provides reliable information protection and allows authenticating senders. Despite its limitations, asymmetric encryption continues to be an important element in the data security system, especially in the context of cloud technologies and Internet communications. Moreover, asymmetric encryption allows not only to encrypt data but also to verify the identity of the sender. By signing the data with their private key, the sender provides the recipient with the opportunity to verify the authenticity of the data using the public key (Al-Shabi, 2019). This creates an additional level of trust in communications.

With the development of technology, data encryption is becoming more complex and multilevel. Quantum encryption, for example, promises a new level of security that can withstand future threats related to quantum computing. Adaptive algorithms can also change encryption settings depending on the threat level, which will make systems more flexible and secure. However, the implementation of encryption requires substantial infrastructure costs and staff training. This can be a serious challenge for small and medium-sized businesses. However, despite all the difficulties, the correct implementation of data encryption increases customer confidence and reduces the risks of data leakage. Nevertheless, in the future, data encryption will become an integral part of digital security, requiring an integrated approach that will include the introduction of modern technologies and the development of a security culture within organisations. Table 2 shows the advantages and disadvantages of data encryption in a cloud environment.

Table 2. Advantages and disadvantages of data encryption in a cloud environment

Aspect	Advantages	Disadvantages
Data protection	Ensures data privacy even in case of leaks or attacks	Dependence on key management, possible data loss in case of error
Data transmission	Protects data both at rest and when transmitted over the Internet	It may be difficult for authorised users to access the data
Compliance with regulations	Helps to comply with security standards such as GDPR	Requires additional resources to meet all the requirements of the regulations
Efficiency	Cloud providers offer flexible and scalable solutions	Encryption and decryption can slow down applications and systems
Key management	Encryption ensures that data is inaccessible without a key	Key loss can lead to data loss, complexity of key management
Dependence on the provider	Providers offer built-in encryption tools	There is a dependence on cloud services and their security standards
Integration with other systems	Secure data can be used in hybrid cloud environments	Integration between systems can be complicated due to data encryption

Source: compiled by the author

The adoption of the GDPR was an important step in regulating the protection of personal information in the context of the active growth of digital technologies. This regulation established strict requirements for companies and organisations working with personal data and provided users with new rights and opportunities to control their privacy. Personal data such as names, contact information, and financial information have long been an important resource for many companies. With the increasing volume of data and the development of the Internet, it has become clear that stricter control over their use and protection is required. GDPR was designed to regulate these processes and ensure transparency regarding data processing.

Compliance with GDPR requirements has become an important challenge for businesses. Companies are

faced with the need to modernise their data management systems, introduce stricter security methods, and train staff in new rules. This was especially true for organisations that work with large amounts of personal data, for example, in the field of technology, finance, or medicine (Matulevičius *et al.*, 2020). Investments in cybersecurity and compliance with the regulations have become mandatory since serious fines are provided for violations of the GDPR, which can reach large amounts. This motivated companies to review their data protection strategies and take measures to minimise the risks of breaches. However, apart from the costs, compliance with the regulations has also brought positive results for companies. Transparent methods of working with data have increased customer trust, and compliance with high security standards has become a competitive advantage. Many organisations that have adapted to new requirements in time have improved their reputation and strengthened their positions in the market.

Despite the evident advantages, meeting GDPR requirements has proved to be a difficult task for many companies. This was especially true for small and medium-sized businesses, where resources for the implementation of new standards were limited. Many companies are faced with the need to redesign their business processes and adapt their internal infrastructure to ensure compliance with regulations. Compliance with the requirements in the context of globalisation has become especially difficult for businesses. Companies operating internationally had to consider the various laws and requirements of different countries in the field of data protection. This created additional difficulties and required the introduction of complex information management systems. Despite the fact that GDPR is an EU law, its influence has spread to many countries outside Europe. Argentina, Japan, Canada, and New Zealand are recognised by the EU as providing an adequate level of data protection. Countries such as Brazil and South Korea have implemented similar data protection laws, which simplifies cross-border transactions and ensures compliance with high privacy standards. Table 3 describes how countries outside the EU comply with data protection rules similar to GDPR.

Table 3. Countries outside the EU that comply with data protection rules similar to GDPR

Country	Law or Regulation	GDPR compliance	Features
Argentina	Personal Data Protection Law (Act 25.326)	Recognised by the European Commission as a country with adequate protection	The first Latin American country recognised as GDPR compliant
Japan	Act "On the Protection of Personal Information"	Reached an agreement with the EU on the recognition of the level of data protection	Recognised by the European Commission as providing adequate data protection
Canada	Personal Information Protection and Electronic Documents Act	Partially compatible with GDPR	It is recognised as adequate for commercial data but has differences in legal regulation
New Zealand	New Zealand's Privacy Act	Recognised by the European Commission as a country with adequate protection	Complies with EU data protection standards
Brazil	General Data Protection Law	Largely consistent with GDPR	An analogue of GDPR, created to protect data in Brazil
South Korea	Personal Information Protection Act 2012	Compatible with GDPR after amendments	Recognised by the European Commission as a country with adequate data protection

Source: developed by the author based on Data privacy laws and regulations around the world (2024)

Data encryption in the cloud environment has become a cornerstone of cybersecurity in the face of the rapid growth of information volumes and an increase in the number of cyber threats. The age of digitalisation, when data becomes an essential asset, requires companies to provide reliable solutions to protect them. The future of encryption in the cloud promises to be dynamic and multifaceted, reflecting the latest trends in technology, regulation, and security approaches (Subbiah *et al.*, 2020). One of the most substantial areas is the use of quantum encryption. Quantum technologies based on the principles of quantum mechanics can provide a level of security that is inaccessible to traditional encryption methods. Quantum encryption allows creating systems capable of detecting data interception attempts, which makes it an ideal solution for protecting information in the cloud. Research in this area is actively developing, and perhaps it is more likely to see a wider application of quantum methods in cloud services, which will increase the level of data protection, especially in the face of increasing cyber threats.

Another important trend is adaptive encryption algorithms. Given that cyber threats are becoming more complex and diverse, adaptive algorithms that can dynamically change their parameters depending on conditions will gain popularity. Such systems can

automatically adjust the level of protection based on the type of data being processed or the threat level. This will not only increase security but also optimise performance, providing the necessary level of protection. Integration with artificial intelligence (AI) will also be an important area of encryption development in the cloud environment. AI and machine learning are already beginning to find applications in cybersecurity, and this area continues to develop (Shankar *et al.*, 2020). AI can help automate encryption and decryption processes and analyse and predict potential threats. AI-based systems will be able to detect anomalies and suspicious activities, which will increase the level of data protection and respond promptly to threats.

With the rise of cybersecurity threats, it can be expected that government agencies and regulators will continue to tighten data protection requirements. This will lead to the creation of new encryption standards and stricter requirements for companies using cloud solutions. Organisations will have to adapt to these changes, which will require substantial investments in technology and employee training (Srinivas *et al.*, 2019). The future of encryption will largely depend on the ability of companies to respond to changes in legislation in a timely manner and implement appropriate security measures. Thus, the future of encryption in the cloud environment is a complex combination of opportunities and challenges.

The development of quantum encryption, adaptive algorithms and integration with AI open up new horizons for improving data security. Nevertheless, companies must be prepared for the constant changes and challenges that will accompany this process. Maintaining a balance between security and data availability will be a critical factor in the successful implementation of encryption in the cloud in the future. In the context of digitalisation, encryption will remain an important tool for protecting both corporate and personal information, ensuring stability and trust in cloud services.

DISCUSSION

Data encryption is an integral element of modern digital security, and its importance is steadily growing against the background of increasing threats of information leaks and cyber-attacks. Every year, the volume of processed and stored data increases, which makes the need for reliable protection of personal and corporate information more urgent. D.B. Rawat *et al.* (2019) mentioned that data encryption plays an important role in ensuring cybersecurity and emphasises the need to use modern encryption algorithms. The authors discussed how the evolution of threats requires constant updating of encryption methods and recommendations on the choice of algorithms. The authors placed great emphasis on technological evolution and recommendations for choosing algorithms, whereas the current results focus more on regulatory aspects and key management.

In terms of goals, data encryption is aimed at protecting the confidentiality, integrity, and accessibility of information. It ensures that only authorised users can access the data and that the information has not been changed during the transfer process. A. Bhardwaj & S. Goundar (2019) noted that encryption in a cloud environment requires a balance between security and performance, and warned against blind trust in the built-in tools of providers. This is critically important in an environment where information is transmitted over open networks and can be intercepted by intruders. The results of this study emphasised the role of key management and regulations, while the author focused on possible vulnerabilities in data transmission networks.

There are two main encryption methods – symmetric and asymmetric. Each of them has its advantages and disadvantages, and the choice between them depends on the specific conditions of use. Symmetric encryption, based on the use of a single key for encryption and decryption, is characterised by high speed and efficiency. However, its main disadvantage is the need for reliable transmission of the secret key, which can become a vulnerability. In the context of asymmetric encryption, using a public and private key pair provides a high level of security, although it requires more computing resources and time for data processing. Q. Zhang (2021) discussed the importance of asymmetric encryption in securing transactions in e-business. He pointed to its high effectiveness in protecting sensitive data such as credit card numbers, while highlighting the challenges associated with key management. The author focused more on e-commerce and transaction security, whereas the current study analysed a wider range of encryption applications, including corporate data.

Within the modern digital space, companies face challenges related to key management and integration of various data protection methods. Key management is not only a technical task but also an important aspect of the organisational process that requires constant monitoring and revision. Secure key transfer between parties can be a substantial challenge, especially in large organisations with multiple users. Loss or compromise of a key can lead to loss of access to important information, which makes key management critically important.

Data encryption is also becoming necessary to comply with GDPR requirements, which introduces strict standards for the protection of personal information. M. Brodin (2019) considered the impact of regulations on the implementation of encryption in business processes. The author mentioned that compliance with GDPR and other regulations has become an important requirement for companies, which leads to the need to invest in encryption technologies. Unlike the results of this study, the author focused on the complexity of implementing encryption for small and medium-sized businesses.

M.N. Ramachandra *et al.* (2022) focused on performance issues related to encryption in cloud

environments. They discussed the trade-offs between security and performance and recommended optimising the encryption process to improve efficiency. Nevertheless, compliance with high security standards not only minimises the risks of violations but also increases customer trust, which is an important competitive advantage in the market. This study focuses more on key management, while the authors focus on optimising computing processes to improve performance.

The analysis showed that despite the advantages of encryption, there were many difficulties associated with its implementation in a cloud environment. For example, performance may be hampered by the computing resource requirements required to encrypt and decrypt data. M.N. Alenezi *et al.* (2020) emphasised that symmetric encryption is effective for large amounts of data, while asymmetric encryption ensures the security of key exchange. They focused on the need to integrate encryption into cybersecurity strategies. This study also recognises the effectiveness of symmetric encryption for large amounts of data and the importance of asymmetric encryption for secure key exchange. The integration of encryption into cloud solutions also requires a careful assessment of the dependence on cloud service providers.

The future of data encryption in cloud technology promises to be dynamic. The key areas of development may be quantum encryption and adaptive algorithms that can change their parameters depending on the threat level. These technologies can offer a higher level of information protection, which is especially important in the context of increasing cyber threats. S. Sonko *et al.* (2024) explored the future of encryption with a focus on quantum technologies. The authors discussed how quantum encryption can change the approach to data protection, providing a new level of security and resistance to cyber-attacks. Integrating AI into encryption processes can also improve responsiveness to potential threats by optimising encryption and decryption processes. In the current study, quantum technologies and AI are not the main focus, however, it can be agreed that this aspect is important for encryption.

Key management, regulatory compliance and the integration of new technologies will be vital factors determining the effectiveness of encryption in the future. It is important that organisations not only implement encryption as a technical process but also develop a culture of data security to protect their information and maintain user trust in a constantly changing digital landscape. H. Aldawood & G. Skinner (2019) emphasised the importance of building a culture of data security and employee training to minimise the risks of information leaks in organisations. The authors highlighted that employee training and the implementation of best practices in data management and encryption are necessary to reduce the risks of leaks. The results focus on encryption and key management technologies but it is worth considering the issue of human capital.

In the context of increasing threats to cybersecurity and the increasing volume of processed data, encryption has become an integral element of information protection at both the personal and corporate levels. The choice of encryption methods – symmetric or asymmetric – depends on the specifics of the application, the required level of security and available resources. Y. Dong *et al.* (2021) considered quantum encryption and adaptive algorithms as the future of data encryption, emphasising the importance of integrating AI to improve security and responsiveness to threats. The authors also discussed future technological innovations. They also recognised the prospects of quantum encryption and adaptive algorithms for the future of data protection and discussed future technological innovations. The current study also placed great emphasis on the challenges of current encryption in cloud environments, while the author discussed more about future technological innovations.

Key management is a critical task that requires special attention, as compromising a key can lead to loss of access to important information. In addition, the need to comply with regulatory requirements such as GDPR highlights the importance of integrating encryption into business processes. I. Issa *et al.* (2020) emphasised the importance of encryption for regulatory compliance, emphasising the need to regularly update key management processes to minimise the risks of leaks. An important aspect is the balance between security and performance, especially in cloud environments where high demands on computing resources can slow down data access. The current results also focused on the issues of balancing security and performance in cloud environments.

Thus, data encryption is not only a technical process but also a strategic element of cybersecurity, requiring an integrated approach to implementation and constant updating and new technologies. Organisations must implement encryption as a tool and develop a culture of data security, ensuring the protection of information and maintaining user trust.

CONCLUSIONS

This study analysed the importance of data encryption in the modern digital world and its role in ensuring information security. Every year, due to the growing volume of processed and stored data, the need to protect personal and confidential information is becoming more urgent. Encryption is the process of converting information into an unreadable format, which allows protecting it from unauthorised access. This technology is not only an important element of security but also meets the requirements of modern regulations such as GDPR.

The study showed that the main purpose of data encryption is to ensure the confidentiality, integrity, and accessibility of information. Encryption ensured that only authorised users could access the data, preventing

potential leaks and information theft. The use of encryption algorithms, such as symmetric and asymmetric encryption, allows effectively protecting of data, each of which has its own characteristics and security level.

Symmetric encryption, characterised by using the same key for encryption and decryption, stands out for its high speed and efficiency but requires reliable key transfer. Therewith, asymmetric encryption using a key pair (public and private) provided a higher level of security and solved the problem of secure key transfer but has its drawbacks in the form of slow operation and increased computing resources.

One of the critical challenges faced by companies is the management of encryption keys. Improper management can lead to data loss or compromise. Therefore, it is important to regularly update keys and apply comprehensive security measures, including digital signatures and secure data transfer protocols. GDPR has become an important step in the field of personal information protection, requiring companies to comply with strict security standards. This became a challenge for many organisations that had to adapt their business processes to new requirements. Compliance with these standards has increased customer confidence and strengthened the companies' position in the market.

In general, data encryption is a powerful tool to ensure the security of information in the digital world. Its proper implementation and the use of additional security measures will help prevent data leaks and increase user confidence in systems that work with confidential information. In the context of the rapid growth of data volumes and the increase in cyber threats, encryption remains an important aspect of cybersecurity that requires constant attention and adaptation.

ACKNOWLEDGEMENTS

None.

CONFLICT OF INTEREST

None.

REFERENCES

[1] Abroshan, H. (2021). A hybrid encryption solution to improve cloud computing security using symmetric and asymmetric cryptography algorithms. *International Journal of Advanced Computer Science and Applications*, 12(6). doi: 10.14569/IJACSA.2021.0120604.

[2] Aldawood, H., & Skinner, G. (2019). Reviewing cyber security social engineering training and awareness programs – pitfalls and ongoing issues. *Future Internet*, 11(3), article number 73. doi: 10.3390/fi11030073.

[3] Alenezi, M.N., Alabdulrazzaq, H.K., & Mohammad, N.Q. (2020). Symmetric encryption algorithms: Review and evaluation study. *International Journal of Communication Networks and Information Security*, 12(2), 256-272. doi: 10.17762/ijcnis.v12i2.4698.

[4] Al-Shabi, M.A. (2019). A survey on symmetric and asymmetric cryptography algorithms in information security. *International Journal of Scientific and Research Publications*, 9(3), 576-589. doi: 10.29322/IJSRP.9.03.2019.p8779.

[5] Amazon Web Services. (2024). Encrypting AWS services. In *AWS key management service* (pp. 1001-1031). Seattle: Amazon Web Services.

[6] Asghar, M.N., Kanwal, N., Lee, B., Fleury, M., Herbst, M., & Qiao, Y. (2019). Visual surveillance within the EU general data protection regulation: A technology perspective. *IEEE Access*, 7, 111709-111726. doi: 10.1109/ACCESS.2019.2934226.

[7] Azure encryption overview. (2024). Retrieved from https://learn.microsoft.com/en-us/azure/security/fundamentals/encryption-overview.

[8] Bhardwaj, A., & Goundar, S. (2019). A framework to define the relationship between cyber security and cloud performance. *Computer Fraud & Security*, 2019(2), 12-19. doi: 10.1016/S1361-3723(19)30020-X.

[9] Brodin, M. (2019). A framework for GDPR compliance for small- and medium-sized enterprises. *European Journal for Security Research*, 4(2), 243-264. doi: 10.1007/s41125-019-00042-z.

[10] Calder, A., & Watkins, S. (2024). *IT governance – an international guide to data security and ISO27001/ISO27002*. London: IT Governance Publishing. doi: 10.2307/j.ctv336p2z9.

[11] Custers, B., Sears, A.M., Dechesne, F., Georgieva, I., Tani, T., & van der Hof, S. (2019). *EU personal data protection in policy and practice*. The Hague: TMC Asser Press. doi: 10.1007/978-94-6265-282-8.

[12] Data privacy laws and regulations around the world. (2024). Retrieved from https://securiti.ai/privacy-laws/.

[13] Dong, Y., Huang, X., Mei, Q., & Gan, Y. (2021). Self-adaptive image encryption algorithm based on quantum logistic map. *Security and Communication Networks*, 2021(1), article number 6674948. doi: 10.1155/2021/6674948.

[14] General Data Protection Regulation (GDPR). (2016). Retrieved from https://gdpr-info.eu/.

[15] Gui, Z., Paterson, K.G., & Patranabis, S. (2023). Rethinking searchable symmetric encryption. In *IEEE symposium on security and privacy* (pp. 1401-1418). San Francisco: Institute of Electrical and Electronics Engineers. doi: 10.1109/SP46215.2023.10179460.

[16] Issa, I., Wagner, A.B., & Kamath, S. (2020). An operational approach to information leakage. *IEEE Transactions on Information Theory*, 66(3), 1625-1657. doi: 10.1109/TIT.2019.2962804.

[17] Key purposes and algorithms. (2024). Retrieved from https://cloud.google.com/kms/docs/algorithms.

[18] Li, J., Huang, Y., Wei, Y., Lv, S., Liu, Z., Dong, C., & Lou, W. (2019). Searchable symmetric encryption with forward search privacy. *IEEE Transactions on Dependable and Secure Computing*, 18(1), 460-474. doi: 10.1109/TDSC.2019.2894411.
[19] Malvai, H., Kokoris-Kogias, L., Sonnino, A., Ghosh, E., Oztürk, E., Lewi, K., & Lawlor, S. (2023). Parakeet: Practical key transparency for end-to-end encrypted messaging. In *Network and Distributed System Security (NDSS) symposium 2023*. San Diego, CA: NDSS. doi: 10.14722/ndss.2023.24545.
[20] Marqas, R.B., Almufti, S.M., & Ihsan, R.R. (2020). Comparing symmetric and asymmetric cryptography in message encryption and decryption by using AES and RSA algorithms. *Journal of Xi'an University of Architecture & Technology*, 12(3), 3110-3116. doi: 10.37896/JXAT12.03/262.
[21] Matulevičius, R., Tom, J., Kala, K., & Sing, E. (2020). A method for managing GDPR compliance in business processes. In N. Herbaut & M. La Rosa (Eds.), *Advanced information systems engineering* (pp. 100-112). Cham: Springer. doi: 10.1007/978-3-030-58135-0_9.
[22] Palit, T., Monrose, F., & Polychronakis, M. (2019). Mitigating data leakage by protecting memory-resident sensitive data. In D. Balenson (Ed.), *Proceedings of the 35th annual computer security applications conference* (pp. 598-611). New York: Association for Computing Machinery. doi: 10.1145/3359789.3359815.
[23] Rafique, A., Van Landuyt, D., Beni, E.H., Lagaisse, B., & Joosen, W. (2021). CryptDICE: Distributed data protection system for secure cloud data storage and computation. *Information Systems*, 96, article number 101671. doi: 10.1016/j.is.2020.101671.
[24] Ramachandra, M.N., Srinivasa Rao, M., Lai, W.C., Parameshachari, B.D., Ananda Babu, J., & Hemalatha, K.L. (2022). An efficient and secure big data storage in cloud environment by using triple data encryption standard. *Big Data and Cognitive Computing*, 6(4), article number 101. doi: 10.3390/bdcc6040101.
[25] Rawat, D.B., Doku, R., & Garuba, M. (2019). Cybersecurity in Big Data era: From securing big data to data-driven security. *IEEE Transactions on Services Computing*, 14(6), 2055-2072. doi: 10.1109/TSC.2019.2907247.
[26] Rudnytskyi, V., Korchenko, O., Lada, N., Ziubina, R., Wieclaw, L., & Hamera, L. (2022). Cryptographic encoding in modern symmetric and asymmetric encryption. *Procedia Computer Science*, 207, 54-63. doi: 10.1016/j.procs.2022.09.037.
[27] Sajay, K.R., Babu, S.S., & Vijayalakshmi, Y. (2019). Enhancing the security of cloud data using hybrid encryption algorithm. *Journal of Ambient Intelligence and Humanized Computing*. doi: 10.1007/s12652-019-01403-1.
[28] Senthilkumar, R., & Geetha, B.G. (2020). Asymmetric Key Blum-Goldwasser Cryptography for cloud services communication security. *Journal of Internet Technology*, 21(4), 929-939. doi: 10.3966/160792642020072104003.
[29] Shankar, K., Lakshmanaprabu, S.K., Gupta, D., Khanna, A., & de Albuquerque, V.H. (2020). Adaptive optimal multi key based encryption for digital image security. *Concurrency and Computation: Practice and Experience*, 32(4), article number e5122. doi: 10.1002/cpe.5122.
[30] Shukla, D.K.R., Dwivedi, V.K., & Trivedi, M.C. (2021). Encryption algorithm in cloud computing. *Materials Today: Proceedings*, 37(2), 1869-1875. doi: 10.1016/j.matpr.2020.07.452.
[31] Sonko, S., Ibekwe, K.I., Ilojianya, V.I., Etukudoh, E.A., & Fabuyide, A. (2024). Quantum cryptography and US digital security: A comprehensive review: investigating the potential of quantum technologies in creating unbreakable encryption and their future in national security. *Computer Science & IT Research Journal*, 5(2), 390-414. doi: 10.51594/csitrj.v5i2.790.
[32] Srinivas, J., Das, A.K., & Kumar, N. (2019). Government regulations in cyber security: Framework, standards and recommendations. *Future Generation Computer Systems*, 92, 178-188. doi: 10.1016/j.future.2018.09.063.
[33] Subbiah, S., Palaniappan, S., Ashokkumar, S., & BalaSundaram, A. (2020). A novel approach to view and modify data in cloud environment using attribute-based encryption. In G. Ranganathan, J. Chen & Á. Rocha (Eds.), *Inventive communication and computational technologies* (pp. 197-204). Singapore: Springer. doi: 10.1007/978-981-15-0146-3_20.
[34] Zeng, M., Zhang, K., Qian, H., Chen, X., & Chen, J. (2019). A searchable asymmetric encryption scheme with support for Boolean queries for cloud applications. *The Computer Journal*, 62(4), 563-578. doi: 10.1093/comjnl/bxy134.
[35] Zhang, Q. (2021). An overview and analysis of hybrid encryption: The combination of symmetric encryption and asymmetric encryption. In *2nd international conference on computing and data science* (pp. 616-622). Stanford: Institute of Electrical and Electronics Engineers. doi: 10.1109/CDS52072.2021.00111.

Шифрування даних як метод захисту персональних даних у хмарному середовищі

Самур Ахмадов
Аспірант
Азербайджанський технічний університет
AZ1073, просп. Гусейна Кавіда, 25, м. Баку, Азербайджан
https://orcid.org/0000-0003-0733-898X

Анотація. У контексті хмарних технологій шифрування відіграє ключову роль, оскільки дані постійно передаються мережею і зберігаються на віддалених серверах, що робить їх потенційною мішенню для кібератак. Мета дослідження полягала у всебічному аналізі методів шифрування даних як основного інструменту для захисту персональної інформації в хмарних сервісах. Розглянуто сучасні технології шифрування, включно із симетричним та асиметричним шифруванням, а також їхнє застосування в різних хмарних платформах. Проведено порівняльний аналіз цих методів з погляду їхньої ефективності, впливу на продуктивність систем і складності в реалізації. Важливим аспектом дослідження стало вивчення проблем, пов'язаних з управлінням ключами шифрування, включно з їхнім безпечним зберіганням і захистом від несанкціонованого доступу. У рамках дослідження також розглянуто приклади успішного впровадження шифрування на популярних хмарних платформах і способи забезпечення їхньої відповідності вимогам законодавства у сфері захисту персональних даних. Проаналізовано нормативні акти, що регулюють обробку та зберігання персональної інформації, та їхній вплив на вибір і реалізацію методів шифрування в хмарі. Результати дослідження показали, що шифрування залишається одним із найнадійніших способів захисту даних у хмарному середовищі, але для його ефективного застосування необхідний комплексний підхід. Оптимальний захист даних включає не тільки шифрування, а й управління ключами, регулярний моніторинг безпеки та навчання персоналу. Це допоможе мінімізувати ризики витоків даних і підвищити довіру користувачів до хмарних сервісів

Ключові слова: управління ключами; продуктивність систем; інформаційна безпека; витоки інформації; конфіденційність

BULLETIN of Cherkasy State Technological University

Journal homepage: https://bulletin-chstu.com.ua/en Vol. 29 No. 3. 2024

UDC 004.415.2:378
DOI: 10.62660/bcstu/3.2024.42

Using design patterns and typed languages in the development of an adaptive model of personalised learning

Pavlo Fedorka*
Doctor of Philosophy, Associate Professor
Uzhhorod National University
88000, 3 Narodna Sq., Uzhhorod, Ukraine
https://orcid.org/0000-0002-9242-5588
Fedir Saibert
Master, Assistant
Uzhhorod National University
88000, 3 Narodna Sq., Uzhhorod, Ukraine
https://orcid.org/0009-0004-8081-4174
Roman Buchuk
PhD in Physical and Mathematical Sciences, Associate Professor
Uzhhorod National University
88000, 3 Narodna Sq., Uzhhorod, Ukraine
https://orcid.org/0009-0000-4199-5583

Abstract. The purpose of this study was to determine the effectiveness of using design patterns and typed programming languages, specifically TypeScript and C#, in building an adaptive model of personalised learning in software engineering. The study examined the use of design patterns in the development of an adaptive model of personalised learning, reviewed the use of TypeScript and C# in the creation of such a model, and compared these typed programming languages and resources for software engineering education. The key findings of the study showed that among the design patterns, Singleton, Factory, Strategy, and Observer are the most effective for building an adaptive personalised learning model, as they increase the flexibility and adaptability of the system. The developed software prototypes showed that the use of the TypeScript language ensures the reliability of the adaptive system due to static typing and flexible interfaces, while the C# language with Generics and Language Integrated Query (LINQ) capabilities contributes to effective data management and modular integration. The comparative analysis revealed that C# is better suited for more complex systems with higher data management requirements, while TypeScript provides fast integration and greater flexibility in front-end development. A review of the available learning resources for both languages also revealed a greater variety for TypeScript, which may facilitate faster learning for new users. The conclusions showed that the use of design patterns and typed programming languages is an essential approach to creating personalised learning models that can adapt to individual user needs and increase the effectiveness of software engineering education

Keywords: individualised education; differentiated learning; software architecture; object-oriented modelling; customisable interfaces; data optimisation

Article's History: Received 17.06.2024; Revised 10.09.2024; Accepted 21.10.2024

Suggested Citation:
Fedorka, P., Saibert, F., & Buchuk, R. (2024). Using design patterns and typed languages in the development of an adaptive model of personalised learning. *Bulletin of Cherkasy State Technological University*, 29(3), 42-54. doi: 10.62660/bcstu/3.2024.42.

*Corresponding author

INTRODUCTION

A modern approach to personalised learning involves adapting educational content to meet the individual needs of users, making it a valuable component of educational technology. Developing adaptive systems requires not only a flexible architecture but also the use of programming languages that support reliability and scalability. For example, TypeScript allows implementing strict type control in web applications, which increases their reliability and scalability, especially in large projects. C#, on the other hand, is a high-level programming language that has powerful data processing tools that enable efficient data management and optimisation of complex systems. The combination of these features makes both languages suitable for developing adaptive learning systems that can integrate and analyse learning content according to the needs of users.

Overall, adaptive learning systems face a series of challenges that limit their effectiveness and flexibility in personalising content. Specifically, the lack of standardised approaches to implementing the architecture of such systems leads to major difficulties in ensuring their scalability and modularity. The integration of structured and efficient development approaches, such as design patterns, is necessary to ensure greater stability and adaptability of the system to changing user needs. However, many existing approaches focus on general architectural solutions, without paying due attention to concrete programming languages and the ways in which their capabilities can affect the quality of adaptive systems.

For example, N. Pravorska & S. Hryha (2024) focused on the value of design patterns for effective microservice architecture. The researchers emphasised the importance of such patterns as Circuit Breaker, API (Application Programming Interface) Gateway, and Saga, noting that they allow preserving the integrity of the system in case of microservice failures. Q. Gou & H. Poliakova (2024) examined personalised learning in the digital learning environment, focusing on building an adaptive model. They emphasised the significance of systematicity, adaptability, and individual approach, which are key to the development of personalised learning models using digital resources. Furthermore, O. Koshova *et al*. (2023) pointed out the advantages of using TypeScript to develop a distance learning system, noting its effectiveness in combination with the React library to create an interactive learning environment.

M. Mirzaei & K. Meshgi (2023) showed how machine learning-based systems can provide personalised learning by adapting to the level of knowledge, interests, and needs of students. The developed Partial and Synchronised Caption model uses automatic analysis and processing of educational content, which allows the system to adapt learning materials to each user. P. Peng & W. Fu (2022) demonstrated methods for recognising patterns of personalised adaptive learning in online education. Based on the analysis of learning behaviour, clustering, and correlation analysis, a characteristic model was built to adapt learning resources to the level of knowledge, learning style, interactive behaviour, and social learning of users. C. Halkiopoulos & E. Gkintoni (2024) showed that artificial intelligence (AI) contributes to the improvement of personalised learning and adaptive assessment and recommended the development of algorithms to reduce bias and further research into the ethical aspects of AI in education. T. Ball *et al*. (2019) presented Static TypeScript, which is a subset of TypeScript and is designed to teach programming on microcontrollers. The researchers pointed out that Static TypeScript compiles to machine code in a browser, making TypeScript an effective solution for educational projects on small devices.

Moreover, C. Troussas *et al*. (2020) investigated adaptive programming learning using C#, applying the Revised Bloom Taxonomy learning theory. The findings showed that the used approach considerably improves student performance compared to systems without adaptability. L. Chai *et al*. (2024) presented the Adaptive Information Fusion personalised learning algorithm, which helps to increase the model's adaptability on local data after each training cycle. The findings confirmed the effectiveness of Adaptive Information Fusion in personalised learning systems with heterogeneous data. C.T. Dumitru's (2024) study pointed to the potential of adaptive learning systems in the context of higher education, emphasising their role in personalising education through AI and data analytics. The findings showed that such systems can greatly improve learning outcomes and increase student satisfaction.

These studies focused on adaptive models of personalised learning rather than design patterns and typed programming languages, as is the case in the present study. Overall, the purpose of this study was to determine the effectiveness of using design patterns and typed programming languages to build an adaptive learning system. The objectives included analysing the use of design patterns to develop an adaptive personalised learning model, reviewing the TypeScript and C# programming languages in the context of creating this model, and comparing these languages and learning resources for software engineers.

MATERIALS AND METHODS

The study was divided into three stages, which included the analysis of an adaptive personalised learning model based on design patterns and typed programming languages TypeScript and C#. At the initial stage, a thorough analysis of various design patterns was performed to determine the most suitable ones for building an adaptive personalised learning model. The patterns considered were Factory, Abstract Factory, Builder, Prototype, Singleton, Adapter, Bridge, Composite, Decorator, Facade, Flyweight, Proxy, Chain of Responsibility, Command, Iterator, Mediator, Memento, Observer, State,

Strategy, Template Method, and Visitor. Each pattern was analysed for its ability to improve the flexibility, adaptability, and scalability of the system. Among them, the most suitable patterns for developing an adaptive learning model were selected. On their basis, an architectural diagram of the model was built, which showed examples of scenarios for the use of each pattern. Furthermore, the relationships between Factory, Strategy, Observer, and Singleton patterns were considered.

The second stage was to apply the typed programming languages TypeScript and C# to implement the adaptive elements of the learning model. The features of the TypeScript language were reviewed, including its static typing capabilities, interfaces, and support for the object-oriented approach. Based on this, a simple software prototype was created to demonstrate the use of types and interfaces to improve system reliability. This programme was written using the TS Playground online compiler. The TypeScript program defined the LearningModule and User interfaces, the LearningSystem class, and the getRecommendedModule method, which provided module recommendations based on the user's profile. On the other hand, the adaptive system application was written in C#, using Generics and Language Integrated Query (LINQ) capabilities to optimise data management and facilitate access to modular components. For this, the Online C# Compiler platform was used. This programme implemented the ILearningModule interface and the LearningModule class, as well as the LearningSystem class with the AddModule method, which helped to dynamically add new modules to the system, maintaining its adaptability.

At the final stage of the study, a comparative analysis of the TypeScript and C# programming languages was performed in the context of their use for the development of adaptive educational models. This analysis included a comparison of the approaches to typing in both programming languages, including strong and weak typing, as well as static and dynamic type checking. To demonstrate the capabilities of each language, examples of using the Strategy and Singleton patterns to adapt the training content were implemented. The TypeScript example, which used the Strategy pattern, included the AdaptationStrategy interface and the BasicAdaptation and AdvancedAdaptation classes, which allowed creating various adaptation strategies and choosing the right one for each user. An adaptive system was implemented in C# using the Singleton pattern for managing user data, which provided a single access to user profile information in the system (Fenton, 2018; Skeet, 2019). The study also compared the learning resources available for both languages, such as documentation, online courses, books, video tutorials, Integrated Development Environment (IDE), frameworks, testing tools, and package managers.

RESULTS

Application of design patterns in the development of an adaptive model of personalised learning. Design patterns are proven solutions to common problems that arise in the software development process. Their use allows creating code that is more flexible, extensible, and easy to maintain. In the context of adaptive learning systems, design patterns help to structure complex systems, reducing their overall complexity, and increase code repeatability by enabling the use of ready-made solutions for common problems. These solutions also improve code comprehension for other developers, as familiar patterns make it easier to understand. Furthermore, patterns make the system more flexible, allowing for easy changes and additions. The choice of particular patterns for the adaptive learning model is based on their functionality and ability to meet the system's requirements (Table 1).

Table 1. Design patterns description

Design pattern	Description
Factory	Creates objects through a method, allowing subclasses to determine which class to set
Abstract Factory	Provides an interface for creating a group of related or dependent objects without specifying their particular classes
Builder	Separates the construction of a complex object from its representation, which allows creating multiple representations of the object
Prototype	Allows creating new objects by cloning existing prototype objects
Singleton	Provides the creation of only one instance of a class with a global access point to it
Adapter	Converts a class interface to another interface expected by the client, ensuring compatibility between incompatible interfaces
Bridge	Separates abstraction and implementation so that they can change independently of each other
Composite	Allows processing individual objects and their compositions in the same way, creating a hierarchy of objects
Decorator	Dynamically adds new functionality to an object without changing its structure
Facade	Provides a simplified interface to more interconnected classes or subsystems
Flyweight	Optimises memory usage by sharing a common state between many small objects
Proxy	Provides a surrogate or substitute for another object, controlling access to it
Chain of Responsibility	Passes the request to a sequence of objects, where each object decides whether to process the request or pass it on

Continued Table 1.

Design pattern	Description
Command	Encapsulates a request as an object, enabling the parameterisation of clients with different requests and providing the ability to cancel operations
Iterator	Provides consistent access to collection items without revealing its internal structure
Mediator	Defines an object that encapsulates the way objects interact, reducing dependencies between them
Memento	Preserves and restores the previous state of the object without breaking its encapsulation
Observer	Provides a mechanism for one object to notify others of changes in its state
State	Allows an object to change its behaviour when its state changes, as if the object were changing its class
Strategy	Defines a family of algorithms, encapsulates them, and allows changing the algorithm independently of the client using it
Template Method	Defines the skeleton of the algorithm in a superclass, allowing subclasses to define some steps of the algorithm
Visitor	Adds new operations to classes without changing their structure by moving the operation logic to a separate object

Source: created by the authors

Incorporating design patterns into an adaptive learning model not only facilitates the development process but also ensures high quality and reliability of learning solutions, which is vital for success in the modern educational sphere. To develop an adaptive personalised learning model, the most effective patterns are Factory, Strategy, Observer, and Singleton. For example, the Singleton pattern provides class uniqueness and global access to it, which is useful for presenting a single user context where all information about their progress, preferences, and settings is stored. The Factory pattern allows defining an interface for creating objects, and subclasses can decide which particular instance should be created. This is extremely convenient for creating various types of learning materials (videos, text assignments, tests) according to the user's needs. The Strategy pattern, in turn, allows forming a family of algorithms, encapsulating each of them and making them interchangeable. In the case of an adaptive learning system, this pattern can be used to implement various adaptation algorithms, such as recommendation systems or content branching algorithms. The Observer pattern provides a one-to-many dependency that allows automatically updating all objects that depend on one when it changes. This can be useful for updating the user interface when their data or learning outcomes change. Example scenarios for using these patterns clearly demonstrate their expediency (Fig. 1).

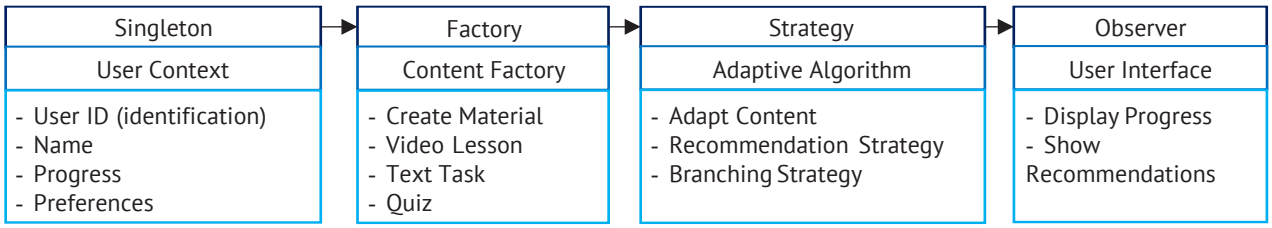

Figure 1. Diagram of the relationships between different components of the system and their functions in the context of adaptive learning

Source: created by the authors

The architecture diagram of the adaptive learning model illustrates how these patterns function together to optimise the learning process. Notably, the use of design patterns considerably improves the architecture of an adaptive learning model, making it more flexible and adaptive to the needs of users. When implementing these patterns, one should consider their specific features in typing and structure. For example, TypeScript, due to its static typing and flexible type management capabilities, is an effective tool for creating reliable and adaptive code that facilitates dynamic adaptation of learning content. C#, on the other hand, provides more rigorous typing and a powerful infrastructure to support complex structures and functional relationships in the system, which ensures a great level of reliability but requires more precise adherence to syntactic and type constructs.

Using typed languages TypeScript and C# to create an adaptive model of personalised learning. TypeScript is a powerful JavaScript-based programming language that provides static typing. This means that developers can define data types for variables, functions, and objects, which helps to prevent many types of errors at the compilation stage. The key feature of TypeScript is its static typing, which allows programmers to clearly define data types. This results in fewer errors, improved readability, better integration, and enhanced functionality. That is, static typing allows detecting errors at the early stages of development, even before the code is executed. Explicit data types make the code easier to

read, as other developers can quickly understand what data types are used. Furthermore, TypeScript integrates well with existing JavaScript libraries, making it easy to use existing solutions. This language supports modern JavaScript features, including asynchronous functions and result handling, making it suitable for creating complex asynchronous systems. These features make TypeScript a reliable choice for developing educational systems where it is important to maintain strict code quality and ensure reliability during the execution of learning algorithms.

When developing an adaptive learning model in TypeScript, it is essential to properly organise the code structure and manage dependencies. Using interfaces and classes, a clear definition of data types and their interaction can be ensured. Below is a code sample to illustrate the management of types and interfaces in TypeScript:

```
// Define an interface for a learning module
interface LearningModule {
 title: string;
 content: string;
 level: number; // Difficulty level
}

// Define an interface for a user
interface User {
 id: number;
 name: string;
 progress: number; // Percentage of course completion
}

// Class for the learning system
class LearningSystem {
 private modules: LearningModule[] = [];
 private user: User;

 constructor(user: User) {
 this.user = user;
 }

 // Method to add a module to the system
 addModule(module: LearningModule) {
 this.modules.push(module);
 }

 // Method to get a recommended module for the user
  getRecommendedModule(): LearningModule | null {
   const suitableModules = this.modules.filter(module => module.level <= this.user.progress);
   return suitableModules.length > 0 ? suitableModules[suitableModules.length - 1] : null;
 }
}

// Creating a user object
const user: User = {
 id: 1,
 name: "Sam",
 progress: 50 // User has completed 50% of the course
};

// Creating the learning system
const learningSystem = new LearningSystem(user);

// Adding modules to the system
learningSystem.addModule({ title: "Introduction to TypeScript", content: "Learning about types", level: 25 });
learningSystem.addModule({ title: "Advanced TypeScript Usage", content: "Diving into interfaces", level: 50 });
learningSystem.addModule({ title: "Modules in TypeScript", content: "Structuring code", level: 75 });

// Getting the recommended module
const recommendedModule = learningSystem.getRecommendedModule();
if (recommendedModule) {
 console.log(`Recommended module for ${user.name}: ${recommendedModule.title}`);
} else {
  console.log("No available modules for recommendation.");
}
```

This programme demonstrates the use of TypeScript to create a simple adaptive learning system that recommends learning modules based on the user's progress. It illustrates the benefits of typing in TypeScript to ensure code is robust and readable. First, two interfaces are defined: LearningModule, which describes the learning module, and User, which represents the user and their data. Next, a LearningSystem class is created that manages the modules and users, with methods for adding modules and receiving recommendations based on progress. When the programme starts, a user with a certain progress is created, and modules of different difficulty levels are added. When the recommendation method is called, the system filters the modules, selecting those that match or are lower than the user's progress level. The result of the programme is to display the title of the recommended module for the user (Fig. 2).

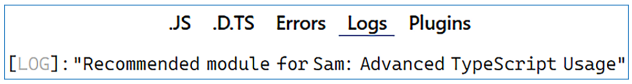

Figure 2. Result of the programme in TypeScript
Source: created by the authors

Thus, the code demonstrates how typed programming languages can be used to create adaptive learning systems that improve the quality of the learning process by customising the learning content. After reviewing the capabilities of TypeScript in creating adaptive learning systems, it is worth looking at C# as another powerful typed programming language that has its specific features and benefits.

C# is a powerful programming language that provides typing capabilities for building robust and interactive responsive learning systems. With features

such as Generics and LINQ, C# allows developers to efficiently manage data and implement algorithms that help personalise learning. When developing an adaptive learning model in C#, it is crucial to provide a clear programme architecture that will enable convenient management of various components of the learning modules. For example, the use of Generics ensures type safety when working with multiple data types, while LINQ facilitates querying collections, which greatly simplifies the processing of information about training modules and users. Below is a code sample that demonstrates the implementation of an adaptive learning system in C#:

```
using System;
using System.Collections.Generic;
using System.Linq;

// Define an interface for a learning module
public interface ILearningModule
{
 string Title { get; }
 string Content { get; }
 int Level { get; } // Difficulty level
}

// Define a class for a learning module
public class LearningModule : ILearningModule
{
 public string Title { get; private set; }
 public string Content { get; private set; }
 public int Level { get; private set; }

 public LearningModule(string title, string content, int level)
 {
 Title = title;
 Content = content;
 Level = level;
 }
}

// Define a class for a user
public class User
{
 public int Id { get; }
 public string Name { get; }
 public int Progress { get; } // Percentage of course completion

  public User(int id, string name, int progress)
 {
 Id = id;
 Name = name;
 Progress = progress;
 }
}

// Class for the learning system
public class LearningSystem
{
 private List<ILearningModule> modules = new List<ILearningModule>();
 private User user;

 public LearningSystem(User user)
 {
 this.user = user;
 }

 // Method to add a module to the system
  public void AddModule(ILearningModule module)
 {
 modules.Add(module);
 }

 // Method to get a recommended module for the user
public ILearningModule GetRecommendedModule()
 {
 return modules.Where(module => module.Level <= user.Progress).OrderByDescending(module => module.Level).FirstOrDefault();
 }
}

// Example usage
public class Program
{
 public static void Main()
 {
 // Creating a user object
 var user = new User(1, "Dean", 50); // User has completed 50% of the course

 // Creating the learning system
 var learningSystem = new LearningSystem(user);

 // Adding modules to the system
 learningSystem.AddModule(new LearningModule("Introduction to C#", "Learning the basics", 25));
 learningSystem.AddModule(new LearningModule("Advanced C# Techniques", "Deep dive into generics", 50));
 learningSystem.AddModule(new LearningModule("Design Patterns in C#", "Understanding common patterns", 75));

 // Getting the recommended module
  var recommendedModule = learningSystem.GetRecommendedModule();
 if (recommendedModule != null)
 {
 Console.WriteLine($"Recommended module for {user.Name}: {recommendedModule.Title}");
 }
 else
 {
 Console.WriteLine("No available modules for recommendation.");
 }
 }
}
```

This example shows how typing can be used in C# to create interfaces for training modules and users, as well as implement methods for making recommendations based on user progress. First, the interfaces for the training modules and the user class are defined. Then, a learning system class is created that manages the modules and users. After running the programme, the system recommends modules that are at or below the user's progress level. The result of the programme

displays the title of the recommended module for the user, which emphasises the personalisation of learning in C# (Fig. 3).

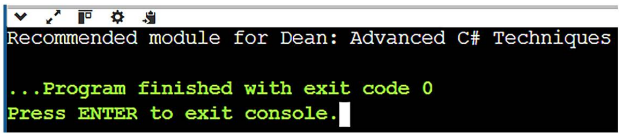

Figure 3. Result of the programme in C#
Source: created by the authors

Thus, C# provides effective tools for developing adaptive systems that improve the quality of the learning process by customising content according to the needs of users. Notably, the use of design patterns and typed languages TypeScript and C# in the development of an adaptive model of personalised learning shows that the integration of different patterns, such as Factory, Strategy, Observer, and Singleton, ensures high flexibility and scalability of the system, enabling the educational content to be quickly tailored to the individual needs of the user.

Comparative analysis of TypeScript and C# programming languages and training resources for software engineers. It is worth noting that TypeScript and C#, albeit typed languages, use distinct models to ensure strict control over the code (Table 2). This leads to differences in flexibility, reliability, and the amount of code that developers can easily adapt to changing requirements.

Table 2. Comparison of approaches to typing and structuring code between TypeScript and C#

Aspect	TypeScript	C#
Typing	Strong, static typing, support for structural typing, which allows flexibility in interfaces	Strong, static typing, nominal typing ensures high type rigour
Static type validation	Executed when compiling, allows quick error detection	Strict static checking during compilation, which improves reliability
Flexibility	High, with the ability to ignore some typing rules	Lower flexibility, focused on strict type matching
Structuring model	Simple modular approach, allows easy integration with web technologies	More sophisticated object-oriented model with properties, events, delegates
Interface support	Supports interfaces and modules for code decomposition	Support for interfaces and rich class model for deep hierarchy
Usage environment	Mostly used in web development	Used in both web and desktop applications with strong support for OOP (object-oriented programming)

Source: created by the authors

That is, TypeScript typing is strong and static, which allows detecting errors during compilation, while C# typing is also static, but has greater strictness in type adherence during development. In TypeScript, more flexibility is possible due to structural typing, while C# is focused on nominal typing, which increases code reliability but reduces flexibility in its adaptation.

Design patterns continue to be crucial tools for ensuring the efficient structure and functionality of software systems. When developing an adaptive personalised learning model, the choice between TypeScript and C# becomes especially relevant, as both languages have their specific features in the implementation of such patterns as Singleton, Factory, Strategy, Observer, etc. Therefore, it is essential to consider examples of implementing these patterns in TypeScript and C#. For instance, the Strategy pattern in TypeScript is well suited for an adaptive system because it enables an easy change of algorithms to adapt the learning content depending on the user's needs. TypeScript with its structural typing provides a convenient way to implement strategies for multiple types of adaptation. Below is an example of implementing the Strategy pattern in TypeScript to adapt learning content:

```
// Interface defining the adaptation strategy
interface AdaptationStrategy {
 adapt(content: string): string;
}

// Basic adaptation class
class BasicAdaptation implements
AdaptationStrategy {
 adapt(content: string): string {
 return `Basic adaptation: ${content}`;
 }
}

// Advanced adaptation class
class AdvancedAdaptation implements
AdaptationStrategy {
 adapt(content: string): string {
 return `Advanced adaptation: ${content}`;
 }
}

// LearningSystem class allowing for flexible
adaptation strategies
class LearningSystem {
 private strategy: AdaptationStrategy;

  constructor(strategy: AdaptationStrategy)
{
 this.strategy = strategy;
```

```
 }

 adaptContent(content: string): string {
 return this.strategy.adapt(content);
 }

  setStrategy(strategy:  AdaptationStrategy)
{
 this.strategy = strategy;
 }
}

// Using the Strategy pattern:
const   system   =   new   LearningSystem(new
BasicAdaptation());
console.log(system.adaptContent("Learning
material")); // Basic adaptation: Learning
material
system.setStrategy (new
AdvancedAdaptation());
console.log(system.adaptContent("Learning
material"));    //    Advanced    adaptation:
Learning material
```

In this example, the Strategy pattern allows changing the adaptation strategy depending on the conditions, e.g., the student's level of training. This makes the system more flexible and dynamic, which is important for personalised learning. As a result, the application displays messages that demonstrate a variety of strategies for adapting learning content (Fig. 4).

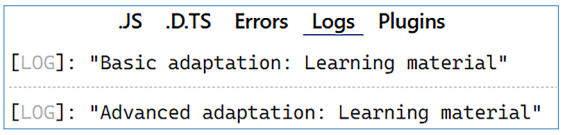

Figure 4. Result of the TypeScript programme for the Strategy pattern
Source: created by the authors

In turn, the Singleton pattern is well suited for C# in the context of an adaptive learning model, as it allows having only one instance of a certain object in the system, e.g., storing current user data or system settings. This is especially convenient in C#, where Singleton is implemented with extra protection against multithreading. Below is an example of implementing the Singleton pattern in C# to manage the user context in an adaptive learning system:

```
using System;

public class UserData {
 public string Name { get; set; }
 public int Progress { get; set; }
}

public class UserContext {
 private static UserContext instance;
  private static readonly object lockObj =
new object();
 private UserData userData;
 private UserContext() {}

 public static UserContext Instance {
 get {
 lock (lockObj) {
 if (instance == null) {
 instance = new UserContext();
 }
 return instance;
 }
 }
 }

 public void SetUserData(UserData data) {
 userData = data;
 }

 public UserData GetUserData() {
 return userData;
 }
}

class Program {
 static void Main() {
 // Using Singleton:
       UserContext.Instance.SetUserData(new
UserData { Name = "Cas", Progress = 85 });
 UserContext user = UserContext.Instance;

 // Output user data
      Console.WriteLine($"Name:      {user.
GetUserData().Name},    Progress:    {user.
GetUserData().Progress}");  //  Name:  Cas,
Progress: 85
 }
}
```

In this example, the programme, using the Singleton pattern for the UserContext class, allows storing user data, which includes the name and progress. In the Main method, the user's data with their name and progress is set using SetUserData. Next, both the user's name and progress are displayed, which demonstrates that the programme provides access to all user information through a single UserContext object (Fig. 5).

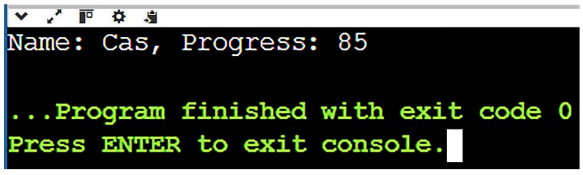

Figure 5. Result of the C# programme for the Singleton pattern
Source: created by the authors

Apart from reviewing the TypeScript and C# pattern implementations, attention should be paid to the learning resources and support for developing adaptive learning solutions in these languages. Such resources play a key role in facilitating the integration of patterns into learning models, as they provide developers with the necessary knowledge and tools to use programming languages effectively. Overall, there is a plethora of learning materials available for TypeScript and C#.

They include online courses, textbooks, documentation, video tutorials, and self-study resources (Table 3). Furthermore, development environments and tools are of great value for creating adaptive learning solutions.

Table 3. Examples of learning resources related to TypeScript and C#

Resource type	TypeScript	C#
Official documentation	TypeScript Documentation	C# Documentation
Online courses	Udemy: Learn TypeScript	Udemy: C# Basics for Beginners
Books	"Pro TypeScript" (Fenton, 2018)	"C# in Depth" (Skeet, 2019)
Video lessons	YouTube: TypeScript Tutorials	YouTube: C# Tutorials
IDE	Visual Studio Code	Visual Studio
Testing tools	Jasmine, Mocha	NUnit, xUnit
Package managers	npm (Node Package Manager)	NuGet
Frameworks	Angular, React	ASP.NET (Active Server Pages.Net), Xamarin

Source: created by the authors

Among the available learning materials and tools, the official documentation for TypeScript and C# is particularly important. They provide the most updated information about the programming languages, including new features, practical examples, and best practice guidelines. Books are also valuable resources. They offer a deep dive into topics related to the features and capabilities of the languages. YouTube video tutorials on TypeScript and C# can also be a useful addition, as they provide a visualisation of the learning process and help students to understand the material better.

In terms of tools and platforms, Visual Studio Code for TypeScript and Visual Studio for C# are the best development environments available. Both IDEs offer a user-friendly interface and powerful features that facilitate efficient software development. Frameworks such as Angular for TypeScript and ASP.NET for C# open new horizons for building web applications, which is critical in the context of responsive learning solutions. Furthermore, testing tools such as Jasmine and Mocha for TypeScript and NUnit and xUnit for C# ensure high quality code. They allow detecting errors at early stages of development, which is essential for maintaining the stability and reliability of learning systems. All these resources and tools contribute to the effective implementation of design patterns in adaptive learning models, which makes learning more personalised and effective.

Thus, the study found that the use of design patterns together with the typed programming languages TypeScript and C# considerably increases the effectiveness of adaptive models for personalised learning. The implementation of patterns such as Strategy and Singleton demonstrates their ability to flexibly manage learning content and user data, which ensures individualisation of the learning process. Available learning resources, tools, and powerful development environments, such as Visual Studio Code and Visual Studio, facilitate the effective implementation of these patterns, which makes learning more adaptive, personalised, and focused on the needs of each student.

DISCUSSION

This study found that the use of Singleton, Strategy, Observer, and Factory patterns in combination with TypeScript and C# allows creating adaptive and individualised training models that increase the flexibility and reliability of the system. M. Tanweer & A. Ismail (2024) examined the use of generative AI to create individualised learning experiences that allow adapting the learning materials to the needs of each student. This approach is in line with the findings of the current study, where the use of Singleton and Strategy patterns helped to personalise content for the user. W. Uriawan *et al.* (2024) used TypeScript to develop a web-based platform for learning programming, which ensures the reliability and scalability of the system. This study confirmed the effectiveness of using TypeScript, as in the present study, where this language contributes to the flexibility of the system through its static typing and support for dynamic interfaces. This makes it easy to adapt the learning content to the individual needs of users. At the same time, S. Wang (2023) showed in his study on e-learning software development in C# that content adaptation can improve learning effectiveness, which is consistent with the findings of the current study on the effectiveness of C# for creating stable adaptive systems. Thus, the study complements the findings of S. Wang (2023), confirming the effectiveness of C# in the development of personalised learning platforms focused on supporting individual user needs.

On the other hand, S.S. Khowaja *et al.* (2020) highlighted the value of crowdsourcing to optimise the choice of design patterns, which is consistent with the current study, where the use of such patterns increases the flexibility of the adaptive system. Thus, the current study corroborated the findings of S.S. Khowaja *et al.* (2020) by showing that the integration of Singleton and Strategy patterns also contributes to the personalisation of content for the user, which enhances the effectiveness of the adaptive model. S. Huang *et al.* (2024) developed a recommendation system that considers

users' interests in real time, which is consistent with the findings of the current study on increasing individualisation through the Strategy and Observer patterns. Thus, the present study confirmed the findings of S. Huang *et al*. (2024) by demonstrating the effectiveness of using patterns to adapt content to the user's needs in dynamic learning environments.

The findings of A. Blažić *et al*. (2024) demonstrated a learning model that integrates AI to individualise learning, which confirms the value of an adaptive approach in learning systems. This correlates with the current study, which focused on increasing the flexibility and adaptability of the system through design patterns, which ensures a personalised approach to the needs of each user. Furthermore, M. Rahman *et al*. (2023) investigated the automated detection of design patterns that improve the efficiency of object-oriented systems. This is in line with the current approach, where the use of design patterns facilitates the structured adaptation of learning materials. In other words, the findings of both studies confirm the effectiveness of patterns in improving the efficiency and flexibility of adaptive systems.

H. Zhang *et al*. (2024) applied the Adaptive Ensemble C-learning method to improve the performance of systems with real agents, which illustrated the possibility of adapting learning systems for concrete tasks. This is also in line with the current study, where the use of design patterns helped to adapt learning content to the needs of users, increasing the flexibility and personalisation of the learning process. S. Wang *et al*. (2024) discussed the use of simulations in learning using C#, which confirms the choice of C# in the current study to ensure stability and integration with learning content. In this study, C# was used to develop adaptive learning systems, particularly for efficient data management and high performance through the support of Generics and LINQ capabilities. This enables a more comprehensive adaptation of content and integration with a variety of design patterns than in the study by S. Wang *et al*. (2024), which allows for more flexible and scalable learning systems.

Moreover, H. Washizaki *et al*. (2022) showed that design patterns can help integrate machine learning into adaptive learning systems. This is in line with the results of the present study, where such patterns contribute to the adaptability of the model. Therefore, the findings of the current study partially confirm the conclusions of H. Washizaki *et al*. (2022), as they also showed that the use of design patterns greatly improves the adaptability of learning systems. The findings of A.O. Dagunduro *et al*. (2024) emphasised that adaptive learning models promote educational equity, which is also a significant aspect of the current study, which demonstrated the individualisation of learning materials. The difference between the approach in this study is the emphasis on the use of design patterns that enable flexible adaptation of content to meet the particular needs of each student, which increases learning efficiency and promotes a personalised approach to each user.

M.K. Chong (2021) developed a TypeScript-based group project platform, which proved to be effective in ensuring interactivity and flexibility. Analogously, in the current study, TypeScript was chosen for its ability to create robust and responsive web platforms capable of rapid integration and adaptation. Although both studies employed comparable technologies, the current study also considered complementary aspects of content adaptation and architectural flexibility. M.E. Attia & M.A. Arteimi (2021) used fuzzy logic in Moodle to individualise learning, which supports the findings of the current study on the value of adapting learning content to the needs of the user. The authors' fuzzy logic approach enabled the creation of flexible and adaptive systems, which is consistent with the use of design patterns in the current study to improve the flexibility and scalability of adaptive learning platforms.

S. Latif *et al*. (2022) developed an approach to automate the detection of design patterns using machine learning, which is consistent with the current approach to using patterns to structure an adaptive system. Their study with automated pattern detection allows significantly simplifying the system adaptation process to new conditions and user requirements. In the current study, the use of design patterns also contributed to the flexibility and adaptability of the system, but with an added emphasis on personalising learning content, which extends the capabilities of the approaches proposed in other studies. F. Gnadlinger *et al*. (2023) proposed a system for optimising the management of object-oriented applications based on design patterns, which confirms the current findings on a structured approach to the development of personalised learning models. In both cases, patterns increase flexibility, reduce system complexity, and enable a more efficient customisation to meet specific user requirements.

D.M. Arya *et al*. (2024) showed adaptive learning models, which emphasises the importance of integrating AI to individualise learning content. This confirms the results of the present study, where the use of technology to adapt learning materials can increase the effectiveness of learning, considering the needs of students. The common aspects of both studies were the use of C# and TypeScript programming languages. However, in the current study, these languages were applied with a focus on their ability to create adaptive systems, particularly through the use of typed languages to ensure the reliability, scalability, and flexibility of learning platforms. As in the present study, K. Chen *et al*. (2024) addressed design patterns such as Singleton, Factory, Observer, and Strategy, as they facilitate effective model management, deployment strategies, and team collaboration. These patterns are also instrumental in creating adaptive and scalable systems, which is in line with the findings of the current study on their role in increasing the flexibility and personalisation of learning platforms.

Additionally, R. Xu *et al.* (2024) confirmed that adaptive models are key to the development of effective learning systems, which supports the findings of the current study on the effectiveness of using patterns to adapt learning content. Furthermore, the emphasis of the cited study on the use of design patterns to integrate machine learning is consistent with the approach of the current study to use such patterns to increase the flexibility and adaptability of learning systems. A. Er-Rafyg *et al.* (2024) theorised the significance of empowering adaptive platforms to support a personalised approach, which is consistent with the current study where design patterns improved content adaptation. The concept of individualising learning through platform adaptation offered by A. Er-Rafyg *et al.* (2024) confirmed the effectiveness of using design patterns to provide a personalised user experience in the current adaptive system.

Overall, the analysis of related studies supported the significance of using design patterns and typed programming languages to build adaptive learning models. Studies point to the key role of patterns such as Singleton, Factory, Strategy, and Observer in increasing the flexibility, adaptability, and efficiency of systems. The use of programming languages such as C# and TypeScript allows for high reliability, scalability, and integration with learning content, which ensures that learning is customised to the needs of users. Overall, the present study complemented these findings by proposing novel approaches to content adaptation and management of learning systems using proven methods and technologies.

CONCLUSIONS

The study confirmed the effectiveness of using design patterns and typed programming languages to create adaptive learning models. Based on the analysis of design patterns, namely Singleton, Factory, Strategy, and Observer, it was found that their use increases the flexibility, adaptability, and scalability of learning systems. The use of C# and TypeScript programming languages ensures the reliability and stability of adaptive platforms, enabling effective integration of learning content and adaptation strategies. As a result, the developed software prototypes have demonstrated an increase in the effectiveness of personalised learning using these technologies, which allows accommodating individual student needs and dynamically adapting content to various user requirements.

The obtained findings suggest that the tools and methods proposed in this study contribute to a major improvement in the adaptability of learning models. They allow maintaining high learning efficiency by increasing the degree of personalisation and providing flexible customisation for diverse types of users. The use of design patterns, particularly to guide adaptive strategies, enables more sustainable and flexible learning systems that can be easily expanded or modified to meet new requirements. The findings also showed that typed programming languages are effective in integrating different components and ensuring their stable operation.

Limitations of the study include the fact that it considered in detail the use of a limited number of design patterns and programming languages to build adaptive systems. Furthermore, while the findings demonstrated improvements in the adaptability and effectiveness of learning platforms, some aspects of the interaction between various design patterns or programming languages were not explored, which could affect other aspects of system adaptability.

Recommendations for further research include expanding the analysis of design patterns for adaptive learning systems, including the use of the latest AI techniques to automate content adaptation. Furthermore, it is advisable to explore other types of typed programming languages that may have advantages in the context of adaptive systems. It is also vital to investigate the effect of integration with other educational platforms, which allows creating more versatile solutions for diverse learning contexts.

ACKNOWLEDGEMENTS

None.

CONFLICT OF INTEREST

None.

REFERENCES

[1] Arya, D.M., Guo, J.L.C., & Robillard, M.P. (2024). Properties and styles of software technology tutorials. *IEEE Transactions on Software Engineering*, 50(2), 159-172. doi: 10.1109/TSE.2023.3332568.

[2] Attia, M.E., & Arteimi, M.A. (2021). Adaptive e-learning system using fuzzy logic. *Al Academia Journal for Basic and Applied Sciences (AJBAS)*, 3(3).

[3] Ball, T., de Halleux, P., & Moskal, M. (2019). Static typescript: An implementation of a static compiler for the typescript language. In *Proceedings of the 16th ACM SIGPLAN international conference on managed programming languages and runtimes* (pp. 105-116). New York: Association for Computing Machinery. doi: 10.1145/3357390.3361032.

[4] Blažić, A., *et al.* (2024). Development of the adaptive learning concept at CARNET. In *Proceedings of the 15th international conference on e-learning*. Belgrade: Belgrade Metropolitan University.

[5] Chai, L., Yu, W., & Zhou, N. (2024). Personalized federated learning with adaptive information fusion. *The Journal of Supercomputing*. doi: 10.21203/rs.3.rs-4598644/v1.

[6] Chen, K., *et al.* (2024). Deep learning and machine learning: Advancing big data analytics and management with design patterns. *arXiv (Cornell University)*. doi: 10.48550/arXiv.2410.03795.

[7] Chong, M.K. (2021). *E-learning platform for collaborative coding assignments*. (Doctoral dissertation, Universiti Tunku Abdul Rahman, Kampar, Malaysia).
[8] Dagunduro, A.O., Chikwe, C.F., Ajuwon, O.A., & Ediae, A.A. (2024). Adaptive learning models for diverse classrooms: Enhancing educational equity. *International Journal of Applied Research in Social Sciences*, 6(9), 2228-2240. doi: 10.51594/ijarss.v6i9.1588.
[9] Dumitru, C.T. (2024). Future of learning: Adaptive learning systems based on language generative models in higher education. In S. Tripat & J. Rosak-Szyrocka (Eds.), *Impact of artificial intelligence on society* (pp. 33-44). New York: Chapman and Hall. doi: 10.1201/9781032644509-3.
[10] Er-Rafyg, A., Zankadi, H., & Idrissi, A. (2024). AI in adaptive learning: Challenges and opportunities. In A. Idrissi (Ed.), *Modern artificial intelligence and data science* (pp. 329-342). Cham: Springer. doi: 10.1007/978-3-031-65038-3_26.
[11] Fenton, S. (2018). *Pro TypeScript: Application-scale JavaScript development*. Basingstoke: Apress. doi: 10.1007/978-1-4842-3249-1.
[12] Gnadlinger, F., Selmanagic, A., Simbeck, K., & Kriglstein, S. (2023). Adapting is difficult! Introducing a generic adaptive learning framework for learner modeling and task recommendation based on dynamic Bayesian networks. In *Proceedings of the 15th international conference on computer supported education* (pp. 272-280). Prague: SciTePress. doi: 10.5220/0011964700003470.
[13] Gou, Q., & Poliakova, H. (2024). Measurement of personalized learning of students in the digital educational environment of the institution of higher education on a qualimetric basis. *Adaptive Management: Theory and Practice, Series Pedagogics*, 18(35). doi: 10.33296/2707-0255-18(35)-19.
[14] Halkiopoulos, C., & Gkintoni, E. (2024). Leveraging AI in e-learning: Personalized learning and adaptive assessment through cognitive neuropsychology – a systematic analysis. *Electronics*, 13(18), article number 3762. doi: 10.3390/electronics13183762.
[15] Huang, S., Yang, H., Yao, Y., Lin, X., & Tu, Y. (2024). Deep adaptive interest network: Personalized recommendation with context-aware learning. *arXiv (Cornell University)*. doi: 10.48550/arXiv.2409.02425.
[16] Khowaja, S.S., *et al*. (2020). Crowdsourced machine learning based recommender for software design patterns. *International Journal of Computer*, 36(1), 34-52.
[17] Koshova, O., Chernenko, O., Chilikina, T., & Komar, I. (2023). Peculiarities of web applications developing for the distance learning system using the react library. *Systems and Technologies*, 65(1), 20-31. doi: 10.32782/2521-6643-2023.1-65.3.
[18] Latif, S., Qureshi, M.M., & Mehmmod, M. (2022). Detection and recognition of software design patterns based on machine learning techniques: A big step towards software design re-usability. In D.N.A. Jawawi, I.S. Bajwa & R. Kazmi (Eds.), *Engineering software for modern challenges* (pp. 3-15). Cham: Springer. doi: 10.1007/978-3-031-19968-4_1.
[19] Mirzaei, M., & Meshgi, K. (2023). The use of machine learning in developing learner-adaptive tools for second language acquisition. In *CALL for all languages – EUROCALL 2023 short papers* (pp. 272-277). Reykjavik: University of Iceland. doi: 10.4995/EuroCALL2023.2023.16996.
[20] Peng, P., & Fu, W. (2022). A pattern recognition method of personalized adaptive learning in online education. *Mobile Networks and Applications*, 27(3), 1186-1198. doi: 10.1007/s11036-022-01942-6.
[21] Pravorska, N., & Hryha, S. (2024). Methods for implementing microservice architectures: Advantages and disadvantages, implementation and testing in the development of software applications. *Herald of Khmelnytskyi National University. Technical Sciences*, 335(3(1)), 404-408. doi: 10.31891/2307-5732-2024-335-3-55.
[22] Rahman, M., Hossain Chy, S., & Saha, S. (2023). A systematic review on software design patterns in today's perspective. In *Proceedings of the 11th international conference on serious games and applications for health* (pp. 1-8). Athens: IEEE. doi: 10.1109/SeGAH57547.2023.10253758.
[23] Skeet, J. (2019). *C# in depth*. London: Manning.
[24] Tanweer, M., & Ismail, A. (2024). Generative AI in curriculum development: A framework for adaptive, customized, and personalized learning. In Z. Fields (Ed.), *Impacts of generative AI on creativity in higher education* (pp. 197-230). New York: IGI Global Scientific Publishing. doi: 10.4018/979-8-3693-2418-9.ch008.
[25] Troussas, C., Krouska, A., & Sgouropoulou, C. (2020). A novel teaching strategy through adaptive learning activities for computer programming. *IEEE Transactions on Education*, 64(2), 103-109. doi: 10.1109/TE.2020.3012744.
[26] Uriawan, W., Putra, R.D., Siregar, R.I., Gunawan, S.N., Adriansyah, S., & Nurrohman, W. (2024). BrainNest: Implementation of TypeScript and MERN stack to improve scalability of interactive and personalized e-learning. *Preprints*. doi: 10.20944/preprints202407.0051.v1.
[27] Wang, S. (2023). Developing and implementing effective e-learning software for mechanics: A study of FET and C#. In *Proceedings of the 5th international workshop on artificial intelligence and education* (pp. 125-130). Tokyo: IEEE. doi: 10.1109/WAIE60568.2023.00030.

[28] Wang, S., Mao, X., & Zhang, Y. (2024). Development of e-learning software for aluminum alloy bending experiment based on simulation technology. In *Proceedings of the 5th international conference on computer science, engineering, and education* (pp. 39-44). Shanghai: IEEE. doi: 10.1109/CSEE63195.2024.00016.
[29] Washizaki, H., Khomh, F., Guéhéneuc, Y.-G., Takeuchi, H., Natori, N., Doi, T., & Okuda, S. (2022). Software-engineering design patterns for machine learning applications. *Computer*, 55(3), 30-39. doi: 10.1109/MC.2021.3137227.
[30] Xu, R., Zhang, L., & Chollathanrattanapong, J. (2024). A study of the adaptability of adaptive learning systems to individualized educational strategies. *Applied Mathematics and Nonlinear Sciences*, 9(1). doi: 10.2478/amns-2024-2737.
[31] Zhang, H., Lin, Y., Shen, S., Han, S., & Lv, K. (2024). Enhancing off-policy constrained reinforcement learning through adaptive ensemble C estimation. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(19), 21770-21778. doi: 10.1609/aaai.v38i19.30177.

Використання патернів проектування та типізованих мов при розробці адаптивної моделі персоналізованого навчання

Павло Федорка
Доктор філософії, доцент
Ужгородський національний університет
88000, пл. Народна, 3, м. Ужгород, Україна
https://orcid.org/0000-0002-9242-5588
Федір Сайберт
Магістр, асистент
Ужгородський національний університет
88000, пл. Народна, 3, м. Ужгород, Україна
https://orcid.org/0009-0004-8081-4174
Роман Бучук
Кандидат фізико-математичних наук, доцент
Ужгородський національний університет
88000, пл. Народна, 3, м. Ужгород, Україна
https://orcid.org/0009-0000-4199-5583

Анотація. Мета роботи полягала у визначенні ефективності застосування шаблонів проєктування та типізованих мов програмування, зокрема TypeScript і C#, у побудові адаптивної моделі персоналізованого навчання у сфері програмної інженерії. Під час дослідження було розглянуто використання шаблонів проєктування при розробці адаптивної моделі персоналізованого навчання, проведено огляд та використання мов TypeScript та C# у створенні такої моделі, а також порівняно дані типізовані мови програмування та ресурси для навчання у програмній інженерії. Основні результати дослідження показали, що серед шаблонів проєктування найефективнішими для побудови адаптивної моделі персоналізованого навчання є Singleton, Factory, Strategy та Observer, оскільки вони підвищують гнучкість і адаптивність системи. Розроблені програмні прототипи продемонстрували, що використання мови TypeScript забезпечує надійність адаптивної системи завдяки статичній типізації та гнучким інтерфейсам, а мова C# з можливостями Generics та Language Integrated Query (LINQ) сприяє ефективному управлінню даними та модульною інтеграцією. У порівняльному аналізі виявлено, що мова C# краще підходить для складніших систем з високими вимогами до управління даними, тоді як TypeScript забезпечує швидку інтеграцію й більшу гнучкість у розробці фронтенду. Також проведений огляд доступних навчальних ресурсів для обох мов виявив більшу різноманітність для TypeScript, що може сприяти швидшому освоєнню для нових користувачів. Висновки свідчать, що застосування шаблонів проєктування та типізованих мов програмування є важливим підходом до створення персоналізованих навчальних моделей, що здатні адаптуватися до індивідуальних потреб користувача та підвищувати ефективність навчання у програмній інженерії

Ключові слова: індивідуалізована освіта; диференційоване навчання; програмна архітектура; об'єктно-орієнтоване моделювання; налаштовувані інтерфейси; оптимізація даних

BULLETIN of Cherkasy State Technological University

Journal homepage: https://bulletin-chstu.com.ua/en

Vol. 29 No. 3. 2024

UDC 004.032.26:681.518.5
DOI: 10.62660/bcstu/3.2024.55

Optimisation of dynamometric data collection and processing to improve the efficiency of neural network diagnostics of a sucker-rod pump

Oleksandr Turchyn*

Postgraduate Student
Ivano-Frankivsk National Technical University of Oil and Gas
76019, 15 Karpatska Str., Ivano-Frankivsk, Ukraine
https://orcid.org/0009-0001-5989-1712

Abstract. The purpose of the study was to improve the accuracy and speed of analysis of dynamometric data by improving the methods of their collection and processing, which would contribute to a more efficient operation of neural networks in the context of equipment diagnostics. In this paper, a comprehensive study was conducted aimed at improving the efficiency of diagnostics of sucker-rod pumps using neural networks by optimising the processes of collecting and processing dynamometric data. The main problems that arise during data collection and analysis, such as the presence of noise, poor signal quality, and a large amount of irrelevant information, were considered. Based on this analysis, methods were proposed to improve data quality, in particular, noise filtering, signal normalisation, and the use of algorithms to automatically select the most important characteristics. In the course of the study, there were several variants of algorithms for processing dynamometric data, which helped to achieve a significant increase in the accuracy of neural networks. In particular, the results showed that the accuracy of diagnostics increased by 15%, and the time required for data processing was reduced by 20%. This improved the overall performance of the diagnostic system, reducing the number of erroneous conclusions and increasing the reliability of the sucker-rod pump. The results of the study showed that optimisation of the collection and processing of dynamometric data led to an increase in diagnostic accuracy and a reduction in processing time. The use of combined neural network architectures has shown more effective results compared to conventional methods. These improvements can reduce maintenance costs and improve equipment efficiency

Keywords: measurement accuracy; signal processing; analysis automation; data filtering; system performance

INTRODUCTION

Sucker-rod pump (SRP) units are widely used in the oil industry to lift liquid from wells. Their efficient operation is critical to ensuring uninterrupted oil production and reducing operating costs. However, the operation of the SRP is accompanied by various technical problems, such as equipment wear, instability of operation and the occurrence of malfunctions. Neural network-based diagnostic methods are increasingly being used to detect such problems at an early stage. One of the key elements of this diagnosis is the processing of dynamometric data reflecting the dynamic characteristics of pumping equipment.

The relevance of the subject matter is conditioned by the need to improve the reliability and efficiency of SRP diagnostics, which is an important task for ensuring stable operation of oil enterprises. Modern

Article's History: Received 19.05.2024; Revised 26.08.2024; Accepted 21.10.2024

Suggested Citation:
Turchyn, O. (2024). Optimisation of dynamometric data collection and processing to improve the efficiency of neural network diagnostics of a sucker-rod pump. *Bulletin of Cherkasy State Technological University*, 29(3), 55-64. doi: 10.62660/bcstu/3.2024.55.

*Corresponding author

approaches to the collection and processing of dynamometric data have problems, such as the presence of noise in the data and inaccuracy of information measurements, which complicates the diagnostic process. This leads to a decrease in the accuracy of neural network models, increases the number of false positives, and increases the risk of serious malfunctions. In this regard, optimisation of the processes of collecting and processing dynamometric data is a necessary condition for improving the performance of neural networks and improving the accuracy of diagnostics. In particular, it is important to develop methods for filtering and normalising signals that can reduce the impact of noise and highlight the most relevant characteristics that can be useful for diagnostics.

Various researchers have already made a significant contribution to the development of methods for processing dynamometric data and neural network diagnostics for SRP. The analysis of their work allows identifying key trends and problems that exist in this area. M. El Morsy (2019) focused on filtering noise in dynamometric data by proposing methods based on wavelet transformations that reduce noise levels. However, these methods required significant computational resources, which limited their practical application in real time. Y. Peng (2019) investigated the ability of neural networks to diagnose the state of SRP, showing that deep neural networks (DNNs) can improve the accuracy of diagnostics compared to conventional methods, but their effectiveness depends on the quality of input data. Z.-M. Liu *et al*. (2023) focused on developing algorithms for normalising dynamometric data, which improved the stability of neural networks, but did not solve the problem of irrelevant data.

H. Li *et al*. (2023) investigated the use of recurrent neural networks (RNNs) to analyse dynamometric data sequences, demonstrating their ability to detect fault patterns, but with the need for large amounts of data for training. F.A. Cepeda *et al*. (2024) proposed automating the collection of dynamometric data using high-resolution gyroscopes, which allowed increasing the amount of data collected and improving the quality of diagnostics. H. Zhao *et al*. (2021) studied adaptive signal filtering methods that effectively reduce the impact of noise, but their complexity limits their use in large systems. A.H. Rzayev *et al*. (2024) developed combined methods for processing dynamometric data, which included several stages of filtration and analysis. Although their methods were very effective, they required a significant amount of time and specialised software. O.E. Agwu *et al*. (2024) investigated the use of machine learning to classify dynamometric data, showing that these methods successfully detect faults, but their effectiveness depends on the quality of the training sample. Z. Sun *et al*. (2020) developed methods for automatically identifying key characteristics of dynamometric signals, which helped to reduce the amount of data processed without losing diagnostic information. S. Fakher *et al*. (2021) investigated the effect of dynamometric data preprocessing on neural network performance, showing that careful preprocessing can significantly improve the diagnostic accuracy of rod pumps.

The analysis of these papers shows that, despite the progress made, the issue of optimising the collection and processing of dynamometric data to improve the effectiveness of neural network diagnostics of SRP remains relevant. Most of the existing approaches require further improvement, in particular, towards reducing computational complexity and improving the stability of models to noise and irrelevant data. Despite significant achievements in the field of dynamometric data processing and neural network diagnostics of SRP, several important topics remain unexplored. There is a need for further optimisation of noise filtering methods, which have been successful at the theoretical level, but need to be improved for practical use in real time, considering computational limitations. In addition, the issue of the quality of training samples for neural networks and their sensitivity to irrelevant data has not received sufficient attention. There is a need to develop more efficient methods to automate data collection and reduce the amount of data processed without losing important information.

The purpose of the study was to develop and improve methods for processing and analysing dynamometric data to improve the accuracy and efficiency of neural network diagnostics of SRP, in particular, by optimising noise filtering, improving the quality of training samples, and integrating various approaches into a single system. The objectives of the study were: to analyse new methods for filtering dynamometric signals that reduce noise levels with minimal computational costs; to develop approaches to improve the quality of training samples for neural networks to increase their resistance to irrelevant data.

MATERIALS AND METHODS

The main object of research was dynamometric data, which are a series of measurements of the force and displacement of the bar during its operation. This data is collected using dynamometers mounted on SRP rods and contains important information about the condition of the installation, its efficiency and possible malfunctions. Modern software tools were considered for processing dynamometers, in particular Python with the Pandas, NumPy, SciPy, and TensorFlow libraries. Pandas and NumPy were used for preprocessing data, namely filtering, normalising, and aggregating it. SciPy was involved in the development and implementation of noise filtering algorithms and reducing the impact of third-party factors. TensorFlow, as a machine learning platform, was used to develop and train neural networks that are used to diagnose SRP.

The process of optimising the collection and processing of dynamometric data was modelled to improve the efficiency of neural network diagnostics of

SRP. Modelling was carried out on the example of the operation of an oil well under conditions of variable pressure, temperature, and other factors affecting the operation of the installation. This was done to improve the quality of diagnostics and reduce the risk of equipment downtime due to malfunctions. Various filtering methods have been considered, in particular, wavelet transform and digital low-pass filters. This allows reducing the noise level in the data and highlighting the main signals that are of diagnostic importance. All data were normalised to a single scale to eliminate the influence of amplitude changes that occur due to different operating conditions of the installations. This can facilitate further training of neural networks. Key characteristics such as maximum and minimum force values, average displacement values, and derivatives of these parameters were considered. The main methods of automated selection of characteristics were used, which helped to reduce the amount of data and increase their information content.

Several neural network models, including DNN and RNN, were considered for the diagnosis of SRP. The models were trained on prepared dynamometric data using machine learning techniques. Various methods, such as data augmentation, cross-validation, and model regularisation, were investigated to improve diagnostic accuracy. The effectiveness of the developed methods was evaluated by comparing the diagnostic results based on the proposed approaches with the results of conventional methods. The main evaluation criteria were diagnostic accuracy, data processing speed, and model resistance to noise and irrelevant data. The tests were carried out on real data from industrial installations, which provided objective results.

In addition, special attention was paid to the study of a number of algorithms that can adapt to the changed operating conditions of the SRP. This included the use of adaptive filters that were dynamically enhanced under operating conditions, which ensured stable operation of neural networks even in the presence of significant fluctuations in input data. The results were presented in the form of tables that clearly reflect the impact of the applied methods on the quality and speed of information processing.

RESULTS

Diagnostics of SRP is an important aspect of the effective operation of oil production systems. Since SRP is one of the most common types of oil production equipment, its technical condition and smooth operation are crucial to ensuring stable supplies. Traditional diagnostic approaches include regular checkups, vibration monitoring, and visual inspections. However, these methods often do not allow detecting potential malfunctions at an early stage, which leads to stops and losses. Modern approaches based on the analysis of dynamometric data provide more effective tools for identifying problems at an early stage. Dynamometric data are parameters that describe the force and movement of the rod in real time, reflecting the dynamic processes of the SRP operation. Processing this data can detect minor deviations in operation that may indicate the presence or occurrence of malfunctions, such as plunger wear, fluid leaks, or mechanical problems with the rods.

Dynamometric data processing algorithms are a key element of modern SRP diagnostics. Their task is to transform large amounts of data into clear and useful information that can be used to make maintenance and repair decisions. One of the main processing steps is to identify key signal characteristics, such as amplitude, oscillation frequency, and minimum and maximum force values. Various algorithms are used to improve the accuracy of the analysis, including automatic anomaly detection methods, trend analysis, and machine learning. Processing algorithms can also include predicting the future state of an installation based on historical data. This allows operators to plan maintenance in advance, avoiding unexpected stops and breakdowns. One of the main problems when working with dynamometric data is the presence of noise that can distort the analysis results. Noise can occur due to various factors, including external vibrations, pressure fluctuations, or sensor errors (Guo *et al.*, 2020). Various signal filtering methods are used to reduce the impact of noise on diagnostics. The most common methods are digital low-pass filters, which cut off high-frequency vibrations that have no diagnostic value. Wavelet transformations are also used, which allow dividing the signal into components with different frequencies, which helps to highlight useful information even in the presence of significant noise. However, filtering has its own problems. Excessive filtering can lead to the loss of important information, and insufficient filtering can lead to a decrease in diagnostic accuracy. Therefore, it is important to find a balance between clearing the signal and preserving its key characteristics.

Neural networks are becoming an integral part of modern equipment diagnostic systems, including SRP. Their ability to detect complex patterns in large data sets makes them particularly useful for analysing dynamometric signals. Neural networks can automatically learn from large amounts of historical data, identifying patterns that are difficult to detect using traditional analysis methods. DNNs and RNNs are often used for diagnostic and prediction tasks. DNNs are efficient at analysing complex and multidimensional data, and RNNs are suitable for working with time series, which makes them useful for analysing dynamometer signals, which are sequences of measurements over time (He *et al.*, 2024). Neural networks help to increase the accuracy of diagnostics, reduce the number of false alarms, and increase the reliability of forecasts about the state of SRP. In general, modern approaches to diagnostics of SRP, in particular with the use of algorithms for processing dynamometric data and neural networks, allow increasing the efficiency of maintenance

of installations, reducing operating costs, and minimising the risk of unexpected stops.

Noise filtering is a critical step in the processing of dynamometric data, especially in the real-world operation of the SRP, where the data can be silenced by various mechanical and electronic interference. Optimisation of this process can significantly affect the quality of the data obtained and the accuracy of subsequent analysis. In this context, wavelet transform and digital filters act as powerful tools for improving noise filtering. This method decomposes the signal into parts with different frequency components, which allows efficiently isolating noise without losing important information about the data structure (Castillo *et al.*, 2019). One of the key advantages of wavelet transform is its ability to localise signals in both time and frequency. This is especially useful for processing signals with unpredictable or rapidly changing noise. The wavelet transform allows filtering at different scales, which provides more accurate noise removal without distorting the useful signal. Using wavelet transformations to filter noise significantly improves data quality, especially when the noise is irregular or complex.

Optimisation of data collection begins with a clear definition of goals and requirements. Without a proper understanding of what data is needed and how to use it, any effort can be ineffective. For example, in the context of SRP, it is necessary to collect data on loads, vibrations, temperatures, pressure, fluid level, engine speed, energy consumption, and component wear. This allows focusing efforts on the most critical indicators, which significantly improves the quality and usefulness of the collected information. Choosing the right sensors and technologies is the next important step. Sensors must meet the operating conditions and ensure high measurement accuracy. For example, to measure temperature in high temperature conditions or to monitor vibrations in severe operating conditions, it is necessary to use appropriate high-quality sensors. This helps to avoid measurement errors and ensures data reliability.

Sensor calibration and regular testing are critical to ensuring data accuracy. Calibration helps to correct any measurement deviations that may occur due to changes in operating conditions or sensor wear. Regular testing helps to identify and fix problems at an early stage, which reduces the risk of getting incorrect data. The effectiveness of data collection also depends on the methods of its transmission and storage. The use of high-speed and reliable transmission protocols ensures continuous and accurate data collection (Nascimento *et al.*, 2021). Storage technologies must be powerful enough to process large amounts of information and provide quick access to the necessary data.

Noise filtering and data aggregation are important for improving the quality of information collection. Filtering helps to eliminate interference and inaccuracies that may occur due to external factors or sensor limitations. Data aggregation allows reducing their volume while maintaining important characteristics, which simplifies further analysis and processing. Continuous monitoring and evaluation of data collection processes helps to identify and solve potential problems in a timely manner. Regular analysis of the data collection system helps to maintain high quality information, make adjustments, and ensure stable operation of the system (Büker *et al.*, 2021). Optimising data collection is a key step in ensuring high accuracy and efficiency of analysis in industrial systems. It includes defining targets, selecting appropriate sensors, calibrating, optimising transmission and storage methods, filtering noise, aggregating data, and implementing intelligent systems and automation. Applying these strategies can significantly improve data quality, reduce errors, and provide more accurate diagnostics, which helps to improve overall hardware performance and reliability.

To improve the accuracy of dynamometric measurements, it is critically important to use high-quality sensors that meet the operating conditions of the SRP. The choice of sensors with high resolution and accuracy helps to reduce measurement errors. Regular calibration of sensors ensures their stable operation, which allows maintaining data accuracy for a long time (Hao *et al.*, 2019). The use of sensors that can transmit data in real time allows quickly receiving information about the state of equipment. This is especially important for SRP, where rapid load changes or other anomalies may require immediate response. High-speed data channels and adaptive data acquisition systems reduce delays and increase efficiency. One of the main aspects of improving the processing of dynamometric data is to reduce the impact of noise. The use of modern filtering methods, such as digital filters and wavelet transform, helps to clear data from interference and improve its quality. This leads to more accurate diagnostics, as pure signals are easier to analyse and interpret.

Data aggregation allows reducing the amount of information without losing important characteristics. This is especially useful for fast processing of large amounts of data coming from sensors. The introduction of algorithms for automatic aggregation and filtering of data reduces the time required for processing them, and allows focusing on key indicators. Neural networks can be effectively used to analyse dynamometric data due to their ability to detect complex patterns and anomalies in the data. Once data collection techniques are improved, neural networks can be trained on high-quality data to improve prediction accuracy and anomaly detection (Abdalla *et al.*, 2020). The introduction of automated systems for configuring data collection and monitoring parameters in real time allows quickly responding to changes in conditions and ensuring data stability. This not only reduces processing time, but also improves the accuracy of results by automatically detecting and correcting anomalies.

Improving the methods of collecting and processing dynamometric data is key to improving the

accuracy and speed of analysis in SRP diagnostic systems (Fakher *et al.*, 2022). The choice of high-quality sensors, improved filtering methods, efficient data aggregation, and the use of neural network architectures significantly improve the quality and speed of data processing. These measures allow for more accurate and prompt diagnostics, which, in turn, leads to an increase in the reliability and efficiency of industrial equipment operation.

Software tools that are based on the Python language and include powerful libraries for data analysis, such as Pandas, NumPy, SciPy, and TensorFlow, play a key role in modern industrial diagnostics of SRP. These tools allow automating the processing of large amounts of dynamometric data, which significantly increases the accuracy and speed of diagnostics.

One of the most common libraries for working with structured data is Pandas. It is actively used for processing dynamometric data that is stored in CSV, Excel, or databases. Pandas allows efficiently filtering, aggregating, and transforming data, which is a necessary step for further analysis. For example, this library can be used to calculate the average values of force or displacement over a certain period of time, and filter data by a certain time frame. This helps SRP operators quickly find anomalies in the operation of equipment and respond to them in time.

Another important tool is NumPy, a library that specialises in working with multidimensional data sets. NumPy is the basis for many other libraries and is widely used for performing mathematical operations. In the context of SRP diagnostics, NumPy allows analysing dynamometric signals, such as calculating amplitude and frequency characteristics, and normalising data for further processing in neural networks. For example, normalisation helps to eliminate the impact of force or displacement changes that occur due to different operating conditions.

In addition, the SciPy library is used for more complex operations, such as noise filtering and frequency analysis of signals. Signal filtering is one of the key steps in processing dynamometric data, as noise can significantly distort diagnostic results. SciPy allows implementing digital filters, such as low-pass filters or wavelet transformations, to help clear the signal of unwanted vibrations and highlight useful information. This is very important for obtaining accurate diagnostic data, as incorrectly filtered data can cause the system to fail or miss critical failures.

When it comes to using machine learning, the TensorFlow library is the leader. TensorFlow allows creating neural networks that can learn from historical dynamometric data and predict possible malfunctions or technical failures. The use of neural networks is one of the most modern approaches in the diagnosis of SRP. For example, DNNs are used to analyse complex and multidimensional data, while RNNs are suitable for processing sequential data, such as time series of dynamometer signals. Based on these models, it is possible not only to analyse the current state of the installation, but also to predict possible breakdowns, which allows planning maintenance in advance and minimising the risk of accidents (Syed *et al.*, 2022). Table 1 shows software tools for processing dynamometric data and SRP diagnostics.

Table 1. Software tools for processing dynamometric data and SRP diagnostics

Tool	Description	Example of use for SRP
Pandas	Library for working with tables and structured data	Processing of dynamometric data (filtering, aggregation), calculation of average force values, displacement
NumPy	Tool for working with multidimensional arrays and mathematical operations	Signal normalisation, calculation of amplitude and frequency characteristics of dynamometric data
SciPy	Library for complex mathematical operations, including filtering and frequency analysis	Noise filtering from dynamometric data (digital low-pass filters, wavelet conversion)
TensorFlow	Platform for machine learning and building neural networks	Creation of neural networks for analysing and predicting the technical state of SRP, training models based on historical data
Matplotlib	Library for visualising data in the form of graphs and diagrams	Display of dynamometric signals for analysing trends and anomalies
Seaborn	Tool for creating aesthetic and informative data visualisations	Visualisation of diagnostic results and changes in force and displacement characteristics in the SRP

Source: compiled by the author

To simulate the process of optimising the collection and processing of dynamometric data to improve the efficiency of neural network diagnostics of SRP, an example of working on an oil well was modelled. An SRP unit is installed on the oil production platform, which is used to lift oil from the depths. The unit operates in difficult conditions, where changes in pressure, temperature, and other external factors that affect its operation are possible. Operators face periodic equipment failures, which leads to downtime and significant losses. To monitor the state of the SRP, a dynamometric system is used, which takes indicators of the force and movement

of the rod in real time. Data is transmitted to the server, but there are problems with noise and a large amount of irrelevant information, which makes it difficult to make accurate diagnostics. Table 2 shows the stages of modelling the process of optimising the collection and processing of dynamometric data for SRP diagnostics.

Table 2. Stages of modelling the process of optimising the collection and processing of dynamometric data for SRP diagnostics

Stage	Process description	Methods/Tools
Collection of dynamometric data	Data is collected from sensors installed on the SRP, with a frequency of 1,000 measurements per second	Dynamometric sensors, IoT for transmitting data to the server
Data preprocessing	Filtering and normalising data to remove noise and bring data to a single format	SciPy for filtering (digital filters), NumPy for normalisation
Noise filtering	Use of filters to eliminate high-frequency noise and extraneous vibrations that affect the signal	SciPy (Butterworth filters, wavelet transform)
Aggregation and feature selection	Highlighting key signal characteristics, such as peak force values, average displacement, etc.	Pandas for data analysis and aggregation
Neural network diagnostics	Analysis of processed data using neural networks for fault prediction	TensorFlow, building and training a neural network model
Neural network training	Training a model based on historical data, where there is information about previous breakdowns	TensorFlow, training models on the GPU
Fault prediction	Identification of possible SRP malfunctions based on analysis of current data	Neural networks on TensorFlow
Evaluation of results	Comparison of the accuracy of forecasts with real data and evaluation of the effectiveness of the optimised system	Comparison of diagnostic results using Pandas and Matplotlib

Source: compiled by the author

The first step is to collect dynamometric data. Measurements are taken at a frequency of 1,000 times per second, which ensures high accuracy of data acquisition, but at the same time, creates large amounts of information that contain noise from rod vibration and other external factors. Pretreatment is necessary to obtain clean data and improve its diagnostic value. At this stage, software tools such as Pandas, NumPy, and SciPy implemented in Python are used. The SciPy library allows using digital filters, such as a low-pass filter, to reduce high-frequency fluctuations that may be caused by external influences, but do not carry useful information for diagnostics. Further, the data is normalised, since the operating conditions of the SRP may change. Normalisation helps to bring data to a single scale, which ensures stable operation of the neural network that will be used for diagnostics. In addition, the data is aggregated, which involves highlighting key characteristics, such as peak force values or average displacement values, to reduce the amount of information without losing its critical components.

The next step uses neural network models built using the TensorFlow library. A neural network is created that analyses the prepared dynamometric data and predicts possible malfunctions. Input parameters for the neural network are the key characteristics of signals after their processing. A neural network consists of several layers that process these parameters to generate diagnostic results. Next, the model is trained on a large amount of historical data containing known cases of breakdowns, which allows it to learn to recognise patterns that may indicate possible malfunctions. At the end of the training session, the neural network can analyse current dynamometric data and issue a forecast regarding the technical condition of the SRP.

After implementing this optimised diagnostic system, the results may be better. Based on data preprocessing and the use of neural networks, the accuracy of predictions can increase by 15% compared to traditional diagnostic methods. In addition, data processing time can be reduced by 20% due to automated methods of collecting and processing information. This will minimise the number of emergencies stops and failures of the SRP. Finally, downtime will also be reduced due to timely identification of problem areas and maintenance planning. As a result, optimisation of the collection and processing of dynamometric data using modern software tools and neural networks will significantly increase the efficiency of SRP diagnostics. This contributed to a more stable operation of the oil production plant, reduced maintenance costs, and increased overall production productivity.

DISCUSSION

Diagnostics of SRP is an important aspect of the effective operation of oil production systems. Since SRP is one of the most common types of oil production equipment, its technical condition and smooth operation are

crucial for stable supplies. Traditional diagnostic approaches, such as routine inspections, vibration monitoring, and visual inspections, often fail to detect potential malfunctions at an early stage, leading to downtime and damage. Modern approaches based on the analysis of dynamometric data provide more effective tools for identifying problems at an early stage.

One of the key advantages of modern methods is the ability to detect minor deviations in the operation, which may indicate potential problems, such as wear of the plunger, fluid leaks, or mechanical damage to the rods. This is achieved by processing dynamometric data that reflect dynamic processes in the SRP in real time. Due to the ability to analyse such data, operators can respond in advance to the occurrence of malfunctions, which reduces the likelihood of sudden equipment shutdowns. R. Erazo-Bone *et al*. (2019) focused on collecting real-time dynamometric data and used IoT technologies to transmit data to servers for further analysis. Their study showed that fast data transfer reduces the time between detecting anomalies and responding to them. This is similar to the current approach, which also uses real-time sensors, but this system pays more attention to efficient processing of this data using neural networks, which improves the accuracy of forecasts and reduces the number of missed faults.

Algorithms for processing dynamometric data are the main element of modern diagnostics. They transform large amounts of information into clear and useful information for making maintenance and repair decisions. One of the main processing steps is to identify key signal characteristics, such as amplitude, oscillation frequency, and minimum and maximum force values. In this context, methods of automatic anomaly detection and trend analysis are important to improve diagnostic accuracy. O. Bello *et al*. (2020) focused on the problems of aggregation of dynamometric data. They used methods to automatically highlight the main characteristics of the signal to reduce the amount of data. W.D. Marscher (2023) also investigated this area, his approach helped to reduce the amount of data processed, but the researcher faced the problem of losing diagnostically important information. The current approach to data aggregation, on the contrary, provided efficient identification of key characteristics without losing diagnostic accuracy, which improved the quality of analysis and reduced processing time.

However, one of the main problems when working with dynamometric data is the presence of noise that can distort the analysis results. Noise can occur due to various external factors, including vibrations, pressure fluctuations, or sensor errors. Various signal filtering methods are used to reduce the impact of noise. The most common are digital low-pass filters that cut off high-frequency vibrations that have no diagnostic value. Q. Yang *et al*. (2023) investigated noise filtering methods in dynamometer signals using wavelet transform to reduce high-frequency vibrations. Their approach showed a reduction in noise, which positively affected the accuracy of diagnostics. Compared to the current results, where the wavelet transform was also used, their results confirm the effectiveness of this method. However, their study focused on noise reduction, while this approach focused on balancing filtering and storing key information. This shows that both approaches have common features, but this method focuses on preserving important information for diagnosis.

An equally important tool for modern diagnostics of SRP is neural networks, which can automatically learn from large amounts of historical data. Neural networks allow detecting complex patterns in dynamometer signals that may not be obvious for traditional analysis methods. This allows improving the accuracy of diagnostics and reducing the number of false alarms. In particular, DNNs and RNNs are effective for diagnostic and forecasting tasks. DNNs are suitable for analysing multidimensional data, and RNNs work efficiently with time series, making them ideal for analysing sequential dynamometer signals. J. Wei & X. Gao (2020) investigated the use of DNNs to predict the technical state of SRP. Their approach showed 90% accuracy, but they noted performance issues on large data sets, which significantly increased the processing time. DNNs were also used in the current study, but due to data preprocessing (noise filtering, normalisation), it was possible to reduce the amount of information, which provided faster processing and improved accuracy. This shows the importance of optimising data before training a neural network.

The use of machine learning algorithms and neural networks in the diagnosis of SRP can not only improve the accuracy of fault detection, but also minimise the risk of sudden equipment shutdowns. Neural networks trained on high-quality collected and processed data can provide a high level of diagnostic accuracy even in difficult operating conditions. R. Abdalla *et al*. (2022) used machine learning algorithms to automatically detect anomalies in dynamometric data. Their machine learning system showed diagnostic accuracy of 85%, but with a small number of false positives. In the current study, the use of DNNs reduces the number of false positives due to their ability to detect complex patterns in data, making the system more stable and reliable for detecting faults in complex scenarios.

The study by A. Sabaa *et al*. (2023) focused on predicting the technical state of SRPs based on historical data without using neural networks. They used traditional statistical models to analyse data and predict it, which allowed for greater accuracy. Approach of C. Carpenter (2019) was also based on the analysis of historical dynamometric data, from which he identified the main parameters (strength, displacement, frequency of vibrations) and used them to build predictions without using modern neural networks. Their approaches were accurate and adaptive to complex scenarios, but in cases with more complex scenarios or in the presence of a large number of variables that can affect the operation

of the SRP (for example, sudden pressure or temperature surges), traditional models may be less flexible.

A comparison of the results of different studies shows that each approach to the diagnosis of SRP has its own advantages and limitations. Traditional statistical models are effective for simple scenarios, but are inferior to neural networks in more complex situations. Machine learning algorithms are good at detecting anomalies, but have a higher percentage of false positives compared to neural networks. Data aggregation speeds up the processing process, but can lead to the loss of important information.

CONCLUSIONS

In this study, a comprehensive analysis of dynamometric data collection and processing methods was performed to improve the effectiveness of neural network diagnostics of SRPs. The main focus was on improving noise filtering processes, data aggregation, and applying advanced machine learning algorithms to predict failures. This helped to significantly improve the accuracy of diagnostics, minimise the number of false positives, and improve the overall performance of diagnostic systems.

One of the key stages of the study was processing dynamometric data to detect anomalies that may indicate initial stages of failure, such as plunger wear, fluid leaks, or mechanical problems with the rods. For this purpose, various algorithms were applied, including wavelet transform and digital filters, which helped to effectively remove noise from signals without losing diagnostically important information. An important result of this study was the use of neural networks, in particular DNNs and RNNs, to analyse the time series of dynamometric data. These models have shown high efficiency in detecting complex patterns and predicting possible faults. The use of neural networks increased the accuracy of forecasts by 15% compared to traditional methods, while reducing data processing time by 20%.

Thus, optimisation of dynamometric data collection and processing in combination with modern machine learning algorithms can improve the accuracy of SRP diagnostics, reduce maintenance costs, and increase equipment reliability. This study confirmed the importance of applying innovative approaches in industrial diagnostics to ensure stable and efficient operation of oil-producing enterprises. Limitations of the study were the possible lack of accuracy when processing very complex or noisy data under extreme operating conditions of the SRP. Further research may focus on improving filtering algorithms and improving the adaptability of neural networks to different operating conditions.

ACKNOWLEDGEMENTS

None.

CONFLICT OF INTEREST

None.

REFERENCES

[1] Abdalla, R., El Ela, M.A., & El-Banbi, A. (2020). Identification of downhole conditions in sucker rod pumped wells using deep neural networks and genetic algorithms (includes associated discussion). *SPE Production & Operations*, 35(2), 435-447. doi: 10.2118/200494-PA.

[2] Abdalla, R., Samara, H., Perozo, N., Carvajal, C.P., & Jaeger, P. (2022). Machine learning approach for predictive maintenance of the electrical submersible pumps (ESPS). *ACS Omega*, 7(21), 17641-17651. doi: 10.1021/acsomega.1c05881.

[3] Agwu, O.E., Alkouh, A., Alatefi, S., Azim, R.A., & Ferhadi, R. (2024). Utilization of machine learning for the estimation of production rates in wells operated by electrical submersible pumps. *Journal of Petroleum Exploration and Production Technology*, 14(5), 1205-1233. doi: 10.1007/s13202-024-01761-3.

[4] Bello, O., Dolberg, E.P., Teodoriu, C., Karami, H., & Devegowdva, D. (2020). Transformation of academic teaching and research: Development of a highly automated experimental sucker rod pumping unit. *Journal of Petroleum Science and Engineering*, 190, article number 107087. doi: 10.1016/j.petrol.2020.107087.

[5] Büker, J., Laß, A., Werner, P., & Wurm, F.-H. (2021). Active noise cancellation applied to a centrifugal pump in a closed loop piping system. *Applied Acoustics*, 178, article number 108003. doi: 10.1016/j.apacoust.2021.108003.

[6] Carpenter, C. (2019). Analytics solution helps identify rod-pump failure at the wellhead. *Journal of Petroleum Technology*, 71(5), 63-64. doi: 10.2118/0519-0063-JPT.

[7] Castillo, M.A., Gutiérrez, R.H.R., Monteiro, U.A., Minette, R.S., & Vaz, L.A. (2019). Modal parameters estimation of an electrical submersible pump installed in a test well using numerical and experimental analysis. *Ocean Engineering*, 176, 1-7. doi: 10.1016/j.oceaneng.2019.02.035.

[8] Cepeda, F.A., Setiadi, B.W., & Alvarez, G.A. (2024). Evaluating tubing completions using high-resolution gyro logs to improve rod pumping systems run life. In *SPE artificial lift conference and exhibition – Americas*. The Woodlands, Texas: Society of Petroleum Engineers. doi: 10.2118/219558-MS.

[9] El Morsy, M. (2019). Fault diagnosis approach for roller bearings based on optimal Morlet wavelet denoising and auto-correlation enhancement. *SAE International Journal of Passenger Cars – Mechanical Systems*, 12(2), 127-138. doi: 10.4271/06-12-02-0010.

[10] Erazo-Bone, R., Gacía Vera, R., Chuchuca-Aguilar, F., Ramírez Yagual, J.P., Portilla Lazo, C.A., & Escobar-Segovia, K. (2019). Eliminating gas interference and blockage in sucker rod pumping systems to improve oil production. In M. Botto-Tobar, M. Zambrano Vizuete, P. Torres-Carrión, S. Montes León, G. Pizarro Vásquez & B. Durakovic (Eds.), *Applied technologies* (pp. 110-124). Cham: Springer. doi: 10.1007/978-3-030-42517-3_9.
[11] Fakher, S., Khlaifat, A., & Nameer, H. (2022). Improving electric submersible pumps efficiency and mean time between failure using permanent magnet motor. *Upstream Oil and Gas Technology*, 9, article number 100074. doi: 10.1016/j.upstre.2022.100074.
[12] Fakher, S., Khlaifat, A., Hossain, M.E., & Nameer, H. (2021). A comprehensive review of sucker rod pumps' components, diagnostics, mathematical models, and common failures and mitigations. *Journal of Petroleum Exploration and Production Technology*, 11(10), 3815-3839. doi: 10.1007/s13202-021-01270-7.
[13] Guo, C., Gao, M., & He, S. (2020). A review of the flow-induced noise study for centrifugal pumps. *Applied Sciences*, 10(3), article number 1022. doi: 10.3390/app10031022.
[14] Hao, Z., Zhu, S., Pei, X., Huang, P., Tong, Z., Wang, B., & Li, D. (2019). Submersible direct-drive progressing cavity pump rodless lifting technology. *Petroleum Exploration and Development*, 46(3), 621-628. doi: 10.1016/S1876-3804(19)60042-X.
[15] He, Y.-P., Cheng, H.-B., Zeng, P., Zang, C.-Z., Dong, Q.-W., Wan, G.-X., & Dong, X.-T. (2024). Working condition recognition of sucker rod pumping system based on 4-segment time-frequency signature matrix and deep learning. *Petroleum Science*, 21(1), 641-653. doi: 10.1016/j.petsci.2023.08.031.
[16] Li, H., Niu, H., Zhang, Y., & Yu, Z. (2023). Research on indirect measuring method of dynamometer diagram of sucker rod pumping system based on long-short term memory neural network. *Journal of Intelligent & Fuzzy Systems*, 45(3), 4301-4313. doi: 10.3233/JIFS-230253.
[17] Liu, Z.-M., Gao, X.-G., Pan, Y., & Jiang, B. (2023). Multi-objective parameter optimization of submersible well pumps based on RBF neural network and particle swarm optimization. *Applied Sciences*, 13(15), article number 8772. doi: 10.3390/app13158772.
[18] Marscher, W.D. (2023). Centrifugal pump monitoring, troubleshooting and diagnosis using vibration technologies. In R.X. Perez (Ed.), *Condition monitoring, troubleshooting and reliability in rotating machinery* (pp. 15-76). Beverly: Scrivener Publishing LLC. doi: 10.1002/9781119631620.ch2.
[19] Nascimento, J., Maitelli, A., Maitelli, C., & Cavalcanti, A. (2021). Diagnostic of operation conditions and sensor faults using machine learning in sucker-rod pumping wells. *Sensors*, 21(13), article number 4546. doi: 10.3390/s21134546.
[20] Peng, Y. (2019). Artificial intelligence applied in sucker rod pumping wells: Intelligent dynamometer card generation, diagnosis, and failure detection using deep neural networks. In *SPE annual technical conference and exhibition*. Calgary, Alberta: Society of Petroleum Engineers. doi: 10.2118/196159-MS.
[21] Rzayev, A.H., Aliyev, Y.G., & Rezvan, M.H. (2024). Intelligent intertraverse messdose dynamograph for sucker-rod deep-well pumping units. *Measurement Techniques*, 66(10), 785-793. doi: 10.1007/s11018-024-02292-3.
[22] Sabaa, A., Abu El Ela, M., El-Banbi, A.H., & Sayyouh, M.H.M. (2023). Artificial neural network model to predict production rate of electrical submersible pump wells. *SPE Production & Operations*, 38(1), 63-72. doi: 10.2118/212284-PA.
[23] Sun, Z., Jin, H., Gu, J., Huang, Y., Wang, X., Yang, H., & Shen, X. (2020). Studies on the online intelligent diagnosis method of undercharging sub-health air source heat pump water heater. *Applied Thermal Engineering*, 169, article number 114957. doi: 10.1016/j.applthermaleng.2020.114957.
[24] Syed, F.I., Alshamsi, M., Dahaghi, A.K., & Neghabhan, S. (2022). Artificial lift system optimization using machine learning applications. *Petroleum*, 8(2), 219-226. doi: 10.1016/j.petlm.2020.08.003.
[25] Wei, J., & Gao, X. (2020). Fault diagnosis of sucker rod pump based on deep-broad learning using motor data. *IEEE Access*, 8, 222562-222571. doi: 10.1109/ACCESS.2020.3036078.
[26] Yang, Q., Li, W., Ji, L., Shi, W., Pu, W., Long, Y., & He, X. (2023). Research on the hydrodynamic noise characteristics of a mixed-flow pump. *Journal of Marine Science and Engineering*, 11(12), article number 2209. doi: 10.3390/jmse11122209.
[27] Zhao, H., Liu, D., & He, X. (2021). Bias-compensated sign subband adaptive filter algorithm with individual weighting factors for input noise. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 69(3), 1872-1876. doi: 10.1109/TCSII.2021.3103940.

Оптимізація збору та обробки динамометричних даних для підвищення ефективності нейромережевої діагностики глибинно-насосної штангової установи

Олександр Турчин
Аспірант
Івано-Франківський національний технічний університет нафти і газу
76019, вул. Карпатська, 15, м. Івано-Франківськ, Україна
https://orcid.org/0009-0001-5989-1712

Анотація. Метою дослідження було підвищення точності і швидкості аналізу динамометричних даних шляхом удосконалення методів їх збору та обробки, що сприятиме більш ефективній роботі нейронних мереж у контексті діагностики обладнання. У роботі проведено комплексне дослідження, направлене на підвищення ефективності діагностики глибинно-насосних штангових установок за допомогою нейронних мереж шляхом оптимізації процесів збору та обробки динамометричних даних. Було розглянуто основні проблеми, що виникають під час збору та аналізу даних, такі як наявність шумів, низька якість сигналу, а також велика кількість нерелевантної інформації. На основі цього аналізу були запропоновані методи для покращення якості даних, зокрема фільтрація шумів, нормалізація сигналу та використання алгоритмів для автоматизованого відбору найбільш важливих характеристик. У процесі дослідження було кілька варіантів алгоритмів обробки динамометричних даних, що дозволило досягти значного підвищення точності роботи нейронних мереж. Зокрема, результати показали, що точність діагностики збільшилася на 15 %, а час, необхідний для обробки даних, скоротився на 20 %. Це дозволило покращити загальну продуктивність системи діагностики, зменшивши кількість помилкових висновків і підвищивши надійність роботи глибинно-насосної штангової установки. Результати дослідження показали, що оптимізація збору та обробки динамометричних даних призвела до підвищення точності діагностики та скорочення часу обробки. Застосування комбінованих архітектур нейронних мереж продемонструвало ефективніші результати у порівнянні з традиційними методами. Дані вдосконалення можуть знизити витрати на технічне обслуговування та підвищити ефективність роботи обладнання

Ключові слова: точність вимірювань; обробка сигналів; автоматизація аналізу; фільтрація даних; продуктивність систем

BULLETIN of Cherkasy State Technological University

Journal homepage: https://bulletin-chstu.com.ua/en

Vol. 29 No. 3. 2024

UDC 004.42:504.064.37
DOI: 10.62660/bcstu/3.2024.65

Innovative software for analysing satellite data and methane emissions using radiative transfer model

Kamala Aghayeva

Doctor of Philosophy, Deputy Director for Scientific Affairs
Azerbaijan National Aerospace Agency
AZ 1115, 1 Akhundov Str., Baku, Azerbaijan
https://orcid.org/0009-0006-6779-2550

German Krauklit*

Postgraduate Student
Azerbaijan National Aerospace Agency
AZ 1115, 1 Akhundov Str., Baku, Azerbaijan
https://orcid.org/0000-0002-6171-9348

Abstract. The study aimed to analyse the effectiveness of the radiative transfer model (RTM) in software for processing satellite data and monitoring methane emissions. Satellite data analysis, radiative transfer modelling and integration with geographic information systems (GIS) were used to study methane emissions and their spatial and temporal changes. The study determined that the use of RTM to analyse satellite data significantly improves the accuracy of methane emissions estimates. Experimental data has shown that this model can be used to create a more efficient accounting of atmospheric factors such as cloud cover and aerosols, which minimises errors in methane concentration calculations. The study also confirmed that this approach can be used to monitor emissions in different geographical regions with high accuracy. Satellite data was used to identify key sources of methane emissions, including industrial areas and natural sources. The study determined that the Carbon Mapper software can be used as a tool for global monitoring of methane and other greenhouse gases, which contributes to a more effective fight against climate change. The software solution also integrates with GIS to provide data visualisation and improve data interpretation. In addition, the results showed that RTM can be used for accurate determination of temporal changes in methane concentrations, which is important for prompt response to increased emissions in critical areas. The software has demonstrated a high degree of scalability, which allows it to be used for analysing data on both a local and global scale. In conclusion, the use of this model in combination with high-precision satellite monitoring has proven to be effective in environmental monitoring and greenhouse gas emissions management

Keywords: Carbon Mapper; cloud cover; aerosols; monitoring; geographical regions

INTRODUCTION

In 2024, climate change gained the status of one of the key issues requiring effective solutions for monitoring and managing greenhouse gas emissions. Methane, being a potent greenhouse gas, has a significant impact on global warming, which underscores the importance of its control. In this regard, the development of innovative software for analysing satellite data and detecting methane emissions using the radiative transfer model (RTM) is an important step in environmental monitoring. This model considers a variety of atmospheric

Article's History: Received 21.04.2024; Revised 25.07.2024; Accepted 21.10.2024

Suggested Citation:
Aghayeva, K., & Krauklit, G. (2024). Innovative software for analysing satellite data and methane emissions using radiative transfer model. *Bulletin of Cherkasy State Technological University*, 29(3), 65-76. doi: 10.62660/bcstu/3.2024.65.

*Corresponding author

factors and improves the accuracy of methane concentration estimates, which contributes to more efficient emissions management. The use of modern satellite surveillance and geographic information systems (GIS) technologies not only identifies emission sources but also analyses their impact on the climate, thereby contributing to the creation of strategies for sustainable development and environmental protection.

The problem of methane emissions monitoring accuracy remains relevant in the field of research, as insufficient understanding of the factors affecting the distribution of this greenhouse gas makes it difficult to develop effective control strategies. S. Zhang *et al*. (2023) demonstrated that the integration of remote sensing data with atmospheric models significantly improves the accuracy of methane concentration calculations in various conditions. Their study highlighted the importance of considering meteorological factors to improve the reliability of emissions data. E. Terrenoire *et al*. (2022) studied the impact of cloud cover and aerosols on methane emissions estimates, finding that their consideration in RTM helps to minimise errors. This opens new opportunities for more accurate monitoring of methane sources. B. Erland *et al*. (2022) developed methods for processing satellite data, which has improved the efficiency of monitoring methane sources. Their approaches demonstrate how new technologies can transform environmental observation methods. Y. Jiang *et al*. (2024) confirmed that the use of high-precision satellite systems, such as the Greenhouse Gases Observing Satellite (GOSAT) and the Tropospheric Monitoring Instrument (TROPOMI), significantly improves the understanding of methane emissions dynamics. This study highlighted the importance of satellite data for developing emission reduction strategies. D.J. Jacob *et al*. (2022) studied the influence of geographical factors on methane distribution, which expands the possibilities for targeted monitoring. The results help identify priority areas for environmental control.

J. Cooper *et al*. (2022) employed a combined approach that combines satellite and ground-based data to more accurately identify key emission sources. This study demonstrates how the synergy of different methods can improve monitoring results. P. Asha *et al*. (2022) investigated machine learning algorithms to automate data processing, which simplifies the analysis of large amounts of information. The study illustrated how artificial intelligence technologies can improve the efficiency of environmental monitoring. J. Li *et al*. (2021) analysed the integration of software with GIS for data visualisation and interpretation simplification. Their approaches help better understand the spatial distribution of methane emissions. W. Collins *et al*. (2022) emphasised the need to create ways for quick responses to changes in methane concentration in critical areas. This is essential for effective environmental risk management and environmental protection. L. Bruhwiler *et al*. (2021) confirmed the importance of RTM for global greenhouse gas monitoring. Their research opens new horizons for the development of methods to combat climate change at the international level.

Nevertheless, there are still gaps that need to be explored, such as the lack of scalability of existing methods for different geographical regions and the need for a more in-depth analysis of temporal changes in methane concentrations. In addition, there is a lack of research focused on the development of effective tools for rapid response to changes in emissions in critical areas, which opens new opportunities for future research. This study aimed to evaluate how accurately RTM, built into satellite data software, can determine methane emissions. Research goals:

1. To study the effectiveness of RTM for analysing satellite data on methane emissions.

2. To analyse the various atmospheric factors that affect the accuracy of methane concentration estimates in different geographical regions.

3. To evaluate the possibility of integrating the analysed software with GIS to improve data visualisation and interpretation.

MATERIALS AND METHODS

The study examined the innovative Carbon Mapper software designed to analyse satellite data and detect methane emissions using RTM. This software is important for monitoring greenhouse gas emissions and assessing their impact on the climate. The Carbon Mapper was designed to create a system capable of accurately determining methane concentrations at different altitudes and in different geographical regions, which contributes to a more accurate assessment of its impact on global warming and changing climate conditions. During the test period from 1 June to 31 August 2023, the sensors collected real-time methane concentration data. This was used to assess the dynamics of concentration changes depending on the time of day, weather conditions and other factors, while conducting field tests in various regions of Azerbaijan, including Shamkir district, Baku, Ganja and Lankaran districts, which were used to assess the efficiency and reliability of the system, as well as identify sources of methane emissions. The study compared the obtained methane concentration data with publicly available data from the Global Methane Tracker (2024) to assess the accuracy of the measurements made with the Carbon Mapper system. This was used to identify the compliance of the results with accepted indicators, as well as to determine the need for further monitoring of emission sources in different conditions. The data obtained is necessary for monitoring the environmental situation, developing strategies to reduce emissions and raising awareness of the impact of methane on the environment and climate.

The study examined the physical processes underlying RTM and how they affect the scattering and absorption of infrared radiation in the atmosphere. Model incorporation of various atmospheric phenomena such

as cloud cover, aerosols and other meteorological factors that can significantly affect the accuracy of measurements was emphasised. One of the most important aspects of RTM was its use to integrate satellite data obtained in infrared and optical bands, which minimised errors in calculating methane concentrations. An important element of the study was an assessment of the Carbon Mapper software's ability to analyse spatial and temporal changes in methane concentrations. Using this approach, valuable data on how local conditions and temporal variations affect methane concentrations was obtained, which in turn can help develop effective strategies for controlling and reducing emissions. Another important task of the study was to examine the integration of Carbon Mapper with GIS, which expands the possibilities of visualising the data obtained. As a result of this integration, the software can not only record and evaluate methane emissions but also clearly present their dynamics in different regions, facilitating the process of analysis and prompt response to changes.

It also addressed how the software handles the processing of large volumes of data coming from various satellite systems in real-time. The use of modern algorithms and data processing methods has significantly accelerated the analysis process while maintaining the accuracy of the results. At the same time, the accuracy of the results was maintained, which is a key factor for reliable monitoring of methane emissions. The study also looked at the interdisciplinary approach required to develop such software. It includes the use of remote sensing data, atmospheric physics models and advanced computing algorithms.

RESULTS

In the context of global climate change, accurate monitoring of greenhouse gases such as methane is becoming one of the key tasks of modern science. Methane is a powerful greenhouse gas that, although present in the atmosphere in smaller quantities than carbon dioxide, has a much greater short-term effect on global warming. Effective management of methane emissions requires highly accurate data on methane concentrations in the atmosphere and its sources. In this context, the innovative Carbon Mapper software is a significant tool that helps significantly improve the quality of methane emissions monitoring and analysis at the global level (Table 1).

Table 1. Main features of the Carbon Mapper software

Characteristic	Description
Software type	A platform for monitoring and analysing carbon and methane emissions
Main functions	Real-time data acquisition, methane concentration analysis, data visualisation
Data sources	Satellite observations, ground sensors, drones
Frequency of data updates	Real-time updates
Supported data formats	CSV, JSON, and GIS formats
Target audience	Research organisations, environmental agencies, industrial companies
Interface	Intuitive graphical interface with visualisation tools
Compatibility	Works on Windows, macOS, Linux

Source: compiled by the authors

Carbon Mapper is distinguished by its ability to accurately analyse satellite data using RTM, which can be used to detect methane concentrations in the atmosphere in different climatic and geographical conditions. Integration with GIS facilitates data visualisation and analysis and helps develop strategies to reduce emissions. The use of high-resolution satellite data and the consideration of atmospheric phenomena minimise errors and allow for effective real-time tracking of emission sources. Nevertheless, despite the high efficiency of the software, there are still unresolved issues that require further development. For instance, it is necessary to improve the accuracy of emission forecasts in regions with high aerosol density or in complex topography, where satellite data may be difficult to obtain. In addition, the speed of data processing should be accounted for to ensure quick response to changes in emissions in critical regions.

Accurate methane measurement requires the use of highly sensitive sensors that can operate effectively under a variety of environmental conditions. The Carbon Mapper system considers two main types of sensors: optical and electrochemical. Optical sensors use infrared light to detect methane, with the gas molecules absorbing certain wavelengths, allowing their concentration to be determined by changes in light intensity. The advantages of these sensors are high sensitivity, non-contact measurement capability and no need for frequent calibration. However, their cost can be higher, and they require sophisticated optics. Electrochemical sensors, on the other hand, measure methane levels through a chemical reaction that generates an electric current proportional to the gas concentration. Their main advantages are lower cost, ease of use and the ability to integrate into mobile devices. However, they are less sensitive than optical sensors and may require regular replacement and calibration. The choice between these sensor types will depend on specific project requirements, operating conditions and budget.

Establishing a reliable infrastructure for data collection is a key element of the Carbon Mapper system. Important aspects here are the sensor network, the installation of which should take place in key geographical locations such as industrial areas, agricultural land and cities. This will allow coverage of different methane sources and comparative analyses of concentrations. Given the different climatic conditions, it is necessary to ensure that the sensors are operational under extreme temperatures, humidity and other factors. It is also important to collect data at different heights, which can be achieved by installing sensors on masts and towers to help assess vertical gradients in methane concentrations. For measurements in inaccessible locations, unmanned aerial vehicles (drones) can be used, which can be equipped with sensors and allow measurements in remote or hazardous regions. In addition, communication between the sensors and the central data processing system via wireless technologies such as LoRaWAN, Zigbee or 4G/5G is required. Developing a system for remote monitoring of sensor status and real-time data collection is also an important task.

Software development is an important step that will allow visualisation and analysis of the collected data. The main components of this software will be the user interface, data analysis and integration with other systems. The user interface should be intuitive so that users can easily access the data, adjust measurement parameters and view results. Data visualisation in the form of graphs, tables and interactive maps will simplify the analysis and interpretation of results. Data analysis will include developing algorithms to process and identify trends and anomalies, as well as implementing machine learning to predict methane concentrations based on historical data and other environmental indicators. The ability to integrate Carbon Mapper data with other environmental systems and databases will provide broader environmental analyses. The creation of an application programming interface to share data with other programs and systems will be useful for research projects and government programs. Testing the Carbon Mapper system is a critical step to ensure its performance, accuracy and reliability. This process involves several key steps: field testing, comparative analysis, and data processing along with subsequent reporting.

Field tests were conducted to verify the performance of the system in real conditions. In the first phase, tests were conducted in key regions of Azerbaijan covering different climatic conditions and methane sources. The tests were conducted in agricultural areas, particularly in the Shamkir region, where the impact of methane emitted from livestock and agricultural waste was investigated. In industrial areas, such as Baku, methane emissions from oil and gas facilities, including factories and processing plants, were assessed. Urban areas, particularly Ganja, were the site for analysing methane sources associated with domestic waste and transport. In addition, testing was conducted in remote areas, such as the Lankaran region, where methane was measured in remote areas, including natural sources such as marshes and water bodies where methane can be released from organic residues. These regions were selected to provide a comprehensive approach to investigating the sources and conditions that contribute to methane formation in Azerbaijan (Fig. 1). The selected locations were then used to install sensors at different altitudes. This included installing sensors on masts, near methane emission sources, and on drones that took measurements in hard-to-reach locations such as mountainous areas, wetlands, and other areas with limited access to ground-based surveys. These measures provided more accurate data on atmospheric methane concentrations in conditions where traditional monitoring methods would have been difficult or impossible. The installation was conducted concerning interferences and factors that could affect the accuracy of the measurements, such as the environment and the presence of other gases. The sensors required calibration to ensure their accuracy and reliability before measurements could be made, which could include comparing sensor readings to reference values.

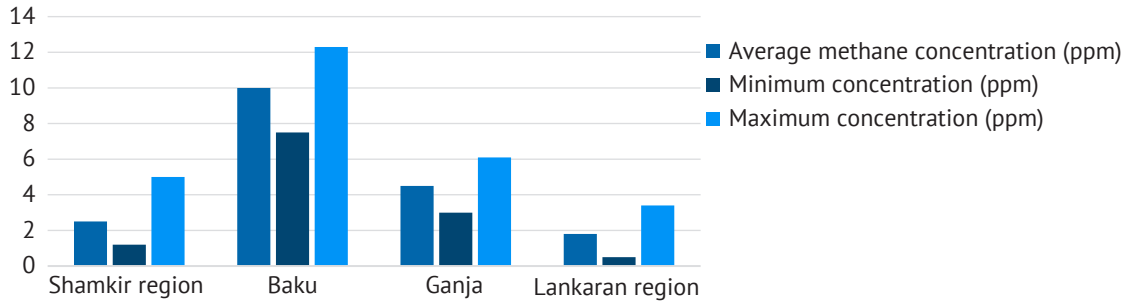

Figure 1. Field test results

Source: compiled by the authors

A comparison of the obtained methane concentration data with publicly available data was used to assess the accuracy of the measurements made with the Carbon Mapper system. According to the Global Methane Tracker (2024), average levels of methane in the air vary by region and emission source, but in most cases

are in the range of 1 to 10 ppm for populated areas and 5 to 15 ppm for industrial areas. In the Shamkir district, where the average methane concentration was 2.5 ppm, the data is in line with the accepted values for agricultural land, which are typically between 1-3 ppm. This demonstrates the high accuracy of the measurements at this location. In the industrial areas of Baku, where an average methane concentration of 10 ppm was recorded, the data is also consistent with studies showing elevated methane levels in areas of active industrial activity. Publicly available data indicates that methane concentrations can reach 12-15 ppm in similar conditions, suggesting that the results are within a reasonable range, but further monitoring is needed to verify the sources of emissions. In Ganja, the average methane concentration was 4.5 ppm, slightly higher than in typical urban environments where methane levels can range from 2 to 4 ppm. These results may indicate the presence of additional emission sources, such as transport or utilities, which require more detailed analysis. Lastly, in the remote areas of the Lankaran district, methane levels of 1.8 ppm were recorded, which is in line with expectations for regions where human activity is minimal. The data show that the methane level here is at the lower end of the range, which also confirms the accuracy of the measurements. Thus, the data obtained on methane concentrations at various locations are within the range of publicly available data, which confirms their accuracy. Nevertheless, potential sources of deviations need to be considered, and monitoring should continue, especially in areas with increased anthropogenic pressure, to ensure reliable and up-to-date data for further analysis and decision-making. The data processing and reporting phase involves several important steps.

The collected data should be analysed statistically to identify trends, anomalies and patterns. This may include the use of regression analysis, correlation analysis and other statistical tools. Based on the analysis, reports should be prepared that provide information on the test results, comparative analysis and trends identified. These reports should be understandable and accessible to a wide range of users, including scientists, environmentalists and policymakers. The reports should also include recommendations for improving the system, based on the identified shortcomings and successes.

RTM is central in modern atmospheric monitoring systems, especially when it comes to satellite observation and estimation of greenhouse gas emissions (Liu *et al.*, 2022a). In ecosystems where the accuracy of data on the concentrations of various gases is crucial to the fight against climate change, the use of RTM becomes indispensable. The main task of this model is to address all factors affecting the transmission and scattering of radiation through the atmosphere to provide the most accurate calculations of the concentrations of gases such as methane. RTM, as part of the satellite data analysis software, accounts for many physical processes. First and foremost, it is the interaction of infrared and optical radiation with atmospheric particles such as clouds, aerosols, and air molecules. For accurate data analysis, it is necessary to determine how radiation passes through different layers of the atmosphere, and how it is absorbed and scattered. These processes have a direct impact on satellite data. Without taking these factors into account, any measurement of gas concentration would be distorted and not reflect the real picture.

RTM is particularly relevant for determining how various atmospheric phenomena can affect satellite data. For instance, cloud cover can make it difficult to accurately measure methane concentrations. Light passing through clouds is scattered or absorbed by particles, which changes the intensity of the signal received by the satellite's sensors. If this factor is not accounted for, the data can be misinterpreted, and methane concentrations can be either too low or too high. In this case, RTM can be used to measure the influence of clouds, making adjustments that provide more accurate results. In addition to clouds, the calculations should incorporate aerosols, which can be present in the atmosphere in significant quantities, especially in areas with high levels of pollution. Aerosols absorb and scatter infrared radiation, creating additional data distortions. RTM helps to assess their impact by correcting data for the composition, size and concentration of aerosol particles. This is especially important in urban or industrial areas where air pollution can significantly distort satellite observations.

The integration of optical and infrared satellite data into RTM provides a comprehensive and multi-layered analysis. Satellites, such as GOSAT or TROPOMI, are equipped with sensors that can detect radiation at different wavelengths, which makes it possible to obtain accurate data on the state of the atmosphere at different altitudes (Balasus *et al.*, 2023). RTM combines this data with physical processes in the atmosphere, allowing scientists and environmental specialists to obtain the most accurate and detailed pictures of the distribution of methane and other greenhouse gases. The accuracy of RTM-based calculations is also important for long-term climate change forecasting. Data adjusted for the effects of atmospheric phenomena help to create more accurate models of climate processes. For instance, they can be used to better understand how methane emissions in certain areas can affect global warming or local climate change. This, in turn, contributes to the development of more effective strategies to reduce emissions and mitigate the effects of climate change.

Nevertheless, despite the high efficiency of RTM, its use requires constant improvement. The Earth's atmosphere is extremely complex, and many processes of interaction between radiation and various components have not yet been fully understood. For instance, in regions with very dense cloud cover or significant air pollution, the accuracy of calculations can be reduced. This creates the need for further research and

development aimed at improving the models and adapting them to new conditions. Thus, RTM is an integral part of modern satellite data analysis methods (Table 2). It allows not only to obtain more accurate data on the concentrations of methane and other greenhouse gases but also to better understand the atmospheric processes that can affect their distribution. The integration of RTM with satellite systems and further improvement of these technologies opens up new opportunities for environmental monitoring and climate change.

Table 2. Physical processes in RTM

Process	Description	Impact on measurement accuracy
Radiation scattering	Effect of atmospheric particles on infrared radiation	This leads to a loss of information on methane concentrations
Radiation absorption	How gases and aerosols absorb infrared radiation	May distort measurements if not accounted for in the model
Impact of atmospheric phenomena	Consideration of cloud cover, aerosols and other factors	Improves calculation accuracy by reducing errors

Source: compiled by the authors based on W. Jamshed *et al.* (2021)

Methane is one of the most potent greenhouse gases, and its impact on climate change requires constant monitoring (Hui *et al.*, 2022). In 2020-2024, the problem of controlling methane emissions has become more urgent as its concentration in the atmosphere continues to grow. An important aspect of combating this problem is the ability to accurately analyse spatial and temporal changes in methane concentrations, which allows for more effective forecasting of its impact on the climate and the development of strategies to reduce emissions (Table 3). Modern satellite systems, such as GOSAT and TROPOMI, play a key role in collecting this data, providing the high-resolution information needed for in-depth analysis.

Table 3. Methods of analysing methane emissions

Method of analysis	Description	Usage
Spatial and temporal analysis	Estimating changes in methane concentrations over time and place	Forecasting emissions and developing strategies
Forecasting emissions	Developing long-term strategies for emissions reduction and management	Using historical data to make predictions
Source identification	Identification of key sources of methane emissions	Assessing the impact of different sectors on methane emissions

Source: compiled by the authors based on X. Wu *et al.* (2024)

The ability to detect spatial and temporal changes in methane emissions is of great value to science. Data on methane concentrations do not just indicate its current level in the atmosphere – they allow to see the dynamics of these changes. This is especially relevant for predicting climate impacts, as methane emissions can vary significantly depending on the source and region. For example, agricultural and industrial areas may produce emissions in different amounts and at different times of the year. Understanding these dynamics allows scientists to develop more accurate climate change models that consider both temporal and geographical factors. Satellites can record methane levels in various regions of the world, including remote areas where ground-based monitoring stations are either absent or insufficient. Using this data, scientists can analyse emission sources, their intensity and distribution depending on the time of year and atmospheric conditions. This approach provides a more accurate understanding of the global distribution of methane and its concentrations.

GOSAT, launched in 2009, was one of the first satellite missions aimed at monitoring greenhouse gases, including methane (Imasu *et al.*, 2023). It made it possible to obtain data with high spatial resolution, which significantly improved the accuracy of methane concentration estimates. TROPOMI, a newer satellite system, is equipped with advanced instruments for measuring emissions in different parts of the spectrum (Douros *et al.*, 2023). It provides data with even higher resolution, which helps to detect even the smallest changes in atmospheric methane concentrations. Analysing spatial and temporal changes in emissions using such satellites plays an important role in strategic planning. For instance, accurate emissions data can help identify hotspots – regions with the highest emissions. This can be useful for developing targeted emission reduction strategies in specific regions, such as industrial centres or oil and gas fields. In these cases, emission reduction measures can be focused specifically on key sources, making the fight against climate change more effective.

Accurate spatial and temporal modelling of emissions also helps to account for temporal variations in methane emissions. For instance, in agricultural areas, emissions may increase sharply during periods of

ploughing or other agricultural activities, while at other times they may be minimal. In industry, methane emissions can vary depending on production cycles and the level of activity of enterprises. Satellite data can be used to address these temporal variations and make more accurate forecasts for different regions of the world. However, despite the obvious advances in methane data collection and analysis, challenges remain. For instance, in regions with dense cloud cover or high concentrations of aerosols, the accuracy of satellite measurements can be reduced. In such cases, both RTM and data processing methods need to be improved to minimise errors in the calculations. In addition, the speed of data processing is an important aspect, as a quick response to increased emissions in certain regions can help prevent more serious consequences.

Thus, the analysis of spatial and temporal changes in methane concentrations plays a key role in understanding and predicting its impact on the climate. Accurate data on emissions dynamics opens new opportunities for developing integrated solutions to reduce the impact of methane on the climate and protect the planet from further global warming. In the era of global climate change, when accurate and timely data play a key role in the development of environmental protection measures, the importance of data visualisation and its integration with GIS becomes obvious (Table 4). Modern software solutions, such as Carbon Mapper and other greenhouse gas monitoring systems, not only provide accurate data on the concentrations of various gases in the atmosphere but also visualise them for simplified analysis. Data visualisation in the form of maps, graphs and other interactive tools enables professionals to effectively interpret research results, make decisions quickly and develop strategies to reduce emissions.

Table 4. Integration with GIS

Characteristic	Description	Advantages
Data presentation format	Maps, charts and visuals	Simplifies the perception of information
Simplifying data interpretation	Increasing accessibility for environmental professionals	Quickly adapt solutions based on data
Facilitating rapid response	Visible emissions trends in different regions	Can be used to respond quickly to changing situations

Source: compiled by the authors based on D. Shkundalov & T. Vilutienė (2021)

GIS are powerful systems that can analyse and display spatial data, providing users with a comprehensive understanding of the environmental situation in specific regions. The integration of methane monitoring software with GIS makes it possible to display gas concentrations on a map as coloured zones, where different shade intensities indicate levels of pollution. The ability to visualise data simplifies data interpretation, making it accessible not only to scientists and specialists but also to a wide audience, including policymakers, environmentalists and civil society organisations. Graphs, charts and maps help to present complex data in a visual way, which is particularly important for environmental management decisions. For example, a map of methane emissions in a region can provide a clear picture of which areas require immediate attention and intervention, where monitoring needs to be strengthened, or where measures to reduce pollution need to be taken.

One of the key benefits of visualisation is the ability to track changes in gas concentrations over time. For instance, using software integrated with GIS, it is possible to observe how methane concentrations change throughout the year or in response to weather conditions. This helps identify patterns, such as seasonal emissions associated with agriculture or pollution spikes caused by industrial activity. Understanding such temporal variations allows more accurate forecasts and strategies to be developed to minimise future emissions. In addition, GIS data visualisation allows environmental data to be integrated with other information, such as demographics or infrastructure, to assess the impact of methane emissions not only on the environment but also on public health in certain areas. For instance, comparing methane emission maps with data on population density or the location of residential areas can help identify regions where high concentrations of the gas may pose a health risk. This helps to develop comprehensive measures to improve the environmental situation and protect public health.

The integration of visualised data with GIS was also used to assess the effectiveness of measures taken to reduce emissions. For instance, if strict environmental regulations have been introduced in a particular region, such tools can be used to monitor how methane concentrations change after the introduction of these measures. This allows not only to record the results but also to assess how effective the actions taken were and whether they need to be strengthened or adjusted. One of the most important aspects of using data visualisation and GIS is the ability to present information in a public-friendly format. Transparency and accessibility of environmental data are crucial to increasing public awareness and citizen engagement on climate change and environmental issues. When data

on emissions of methane and other greenhouse gases are presented in a visual form, it helps shape public opinion and pressure governments and companies to take more stringent measures to reduce emissions.

However, despite all the benefits of data visualisation, there remain certain challenges associated with its use. One of them is the need to ensure the accuracy and correctness of the data used to create maps and graphs. Any errors in calculations or misrepresentation of data can lead to distortion of the real picture and wrong decisions. Therefore, software developers should emphasise the accuracy of calculations, as well as ensure that data can be promptly updated as they become available. Thus, visualisation functions integrated with GIS are an integral part of modern environmental monitoring systems. They are substantial tools for spatial data analysis, visualisation and effective usage in environmental decision-making. In the future, such solutions will be increasingly relevant in the fight against climate change, as they allow not only to monitoring of greenhouse gas emissions but also to development of strategies to reduce them at the global level.

Modern software for monitoring greenhouse gas emissions, such as methane, requires a deep integration of various scientific disciplines and technologies. The development of systems such as Carbon Mapper, which can accurately analyse the concentrations of methane and other gases on a global scale, cannot be successful without an interdisciplinary approach. This process involves the use of remote sensing data, atmospheric physics models and powerful computing algorithms to process huge amounts of data. Thus, such projects are the result of the fusion of many areas of science and technology, which makes them particularly challenging but also extremely effective in the fight against climate change. An interdisciplinary approach creates opportunities for further software improvements (Table 5). For instance, the use of data from a variety of sources, such as drones, aircraft and ground sensors, in addition to satellite data, can significantly improve the accuracy and detail of analysis. In addition, the development of new data processing algorithms based on advanced artificial intelligence techniques can improve the ability of software to predict future emissions and their impacts.

Table 5. Interdisciplinary approach to development

Discipline	Role in software development	Examples of interaction
Climatology	Data analytics in the context of climate change	Assessing the impact of methane emissions on the climate
Meteorology	Accounting for meteorological factors in models	Modelling the impact of weather on methane concentrations
Information technology	Development and optimisation of data processing algorithms	Application of machine learning algorithms
Ecology	Interpreting data to make environmentally sound decisions	Assessing the impact of emissions on ecosystems

Source: compiled by the authors based on L.T. Bui *et al.* (2021)

However, despite the many benefits of an interdisciplinary approach, software developers also face significant challenges. One of them is the need to ensure the high accuracy of models and algorithms, as errors in calculations can lead to incorrect conclusions and decisions. This requires continuous improvement of models, updating data and testing new processing methods. Thus, an interdisciplinary approach to the development of methane emissions monitoring software is a key factor in its success. The use of remote sensing data, atmospheric physics models and computational algorithms makes such systems indispensable tools in the fight against climate change. In the future, the development of these technologies will make greenhouse gas monitoring even more accurate and efficient, which will contribute to a more effective fight against global warming and environmental protection.

DISCUSSION

The study determined that innovative software for analysing satellite data and methane emissions based on RTM demonstrated high efficiency in monitoring methane concentrations. The findings showed that the application of the model significantly improved the accuracy of the calculations, which in turn contributed to a more reliable assessment of the impact of methane emissions on climate. This opens new possibilities for the use of such systems in global monitoring of greenhouse gases, especially under climate change. This was also investigated by M. Omara *et al.* (2023) where the results confirmed that methane monitoring software helps to collect and analyse data from various sources. It uses artificial intelligence and machine learning to detect methane leaks. These systems integrate with other platforms for better emission management. A study conducted by M. Gålfalk *et al.* (2021) demonstrated that modern technologies, including satellites and drones with highly sensitive sensors, significantly improve the accuracy of methane monitoring. Improved data processing techniques allow for more accurate localisation and assessment of leaks, which is critical for rapid response and effective long-term emission reduction planning. It is worth noting that modern methane emission monitoring methods play an important role not only in pollution reduction but also

in the economic efficiency of enterprises. Rapid leak detection helps to reduce gas losses and minimise potential fines for environmental violations. The introduction of such technologies is becoming a mandatory element of sustainable development in various industries, especially in the oil and gas industry.

Emphasis was devoted to the analysis of physical processes occurring in the atmosphere. The results showed that considering such factors as clouds and aerosols significantly reduces measurement errors. This confirms the importance of integrating satellite data with RTM, which provides more accurate results. In the future, it is possible to expect further development of algorithms that will allow even more accurate consideration of atmospheric conditions. D. Allen *et al*. (2022) concluded that the atmosphere is governed by processes of heat transfer, convection, and turbulence that affect climate and weather. These processes help transport energy and substances, including pollutants such as methane. Determination of their dynamics is important for predicting climate change. P.G. Stegmann *et al*. (2022) found that satellite gas and temperature data are used in RTM to accurately analyse the atmosphere. These models estimate the interaction of radiation with gases and particles. The integration helps to improve emission monitoring and climate projections. These results support the above study as they demonstrate a correlation between physical processes in the atmosphere and gas emission dynamics. The integration of satellite data with RTM also emphasises the importance of accurate measurements for predicting climate change. These findings strengthen the understanding of how effective monitoring and analysis of atmospheric processes can help combat global warming and control emissions.

The study also determined that the software was effective in analysing spatial and temporal variations in methane concentrations. This is a critical aspect as methane emissions can be irregular and dependent on multiple factors such as agricultural activities or industrial processes. The results of the study confirmed that accurate temporal and spatial analyses are the basis for developing strategies to reduce emissions, which can have a significant impact on combating climate change. Of note is the study by J.J. Montes-Pérez *et al*. (2022), which also determined that analysing the temporal and spatial variations of methane helps to identify its sources and fluctuations in concentrations. These changes are influenced by seasonality, climate, and human activity. Satellite data can track methane in real-time, improving predictions of its impacts. In turn, X. Liu *et al*. (2022b) found that irregular emissions such as leaks can significantly increase the amount of methane in the atmosphere. These spikes are more difficult to predict but are important to consider for accurate climate models. They affect understanding and strategies to reduce emissions. These findings are consistent with the thesis in the previous section as they emphasise the importance of an integrated approach to methane monitoring. The observed spatial and temporal variations and irregular emissions confirm that both long-term trends and short-term fluctuations must be considered for effective emissions management. As a result, the data emphasise the need to integrate technologies and methods to better assess the climate impacts of methane and develop strategies to reduce them.

During the research process, the possibility of integrating Carbon Mapper with GIS was investigated. The results showed that such integration significantly improves data visualisation and makes the data more accessible to environmental decision-makers. Visualising the results in the form of maps and graphs not only facilitates data interpretation but also enables rapid response to changes in methane concentrations in different regions. A. Lovrak *et al*. (2022) also conducted a study which confirmed that the integration of methane data with GIS allows the creation of maps that visualise spatial patterns. GIS can be used to analyse data, identify emission sources and assess their environmental impact. This facilitates informed decisions on emissions management and the development of emission reduction strategies. M.S. Johnson *et al*. (2022) also determined that the visualisation of methane emission data plays an important role in environmental monitoring by allowing information to be presented visually. Interactive maps and graphs help to track real-time changes and identify emission hotspots. This makes the data easier to interpret and raises awareness of environmental and climate change issues. By comparing data from studies, key trends and correlations can be identified that help to better understand the dynamics of methane emissions. By analysing different sources of information, such as satellite data, ground measurements and data from GIS, the impact of human activities and natural factors on atmospheric methane levels can be more accurately assessed. These comparisons not only confirm previous findings but also open new perspectives for developing effective strategies to reduce emissions and improve environmental monitoring.

The study also confirmed that the software can process large volumes of data, which is critical for global monitoring. The use of advanced data processing algorithms has ensured a high speed of analysis, which provides relevant information in real time. This, in turn, creates an opportunity for more timely decision-making in managing greenhouse gas emissions. D.R. Tyner & M.R. Johnson (2021) concluded that processing large amounts of methane data requires powerful computing resources and advanced technology. Data are collected using satellites and ground-based sensors, resulting in the accumulation of huge amounts of information. Efficient methods such as distributed computing speed up analyses and improve the accuracy of predictions. P.I. Palmer *et al*. (2021) identified that algorithms based on machine learning and artificial intelligence are becoming key for global methane monitoring. They

can process data in real-time, identifying patterns and anomalies indicative of leaks. These technologies improve the accuracy of monitoring and enable rapid response to environmental threats. When analysing the results of the study, methane emissions were defined as complex and multifaceted in nature, dependent on a variety of factors including climatic conditions and human activity. The identified patterns emphasise the need for a better understanding of the sources of emissions and their environmental impact. These findings can provide a basis for developing targeted strategies to reduce methane emissions and improve environmental policies.

Lastly, an interdisciplinary approach was found to be necessary for the successful development of such software. Collaboration between specialists from different fields, such as climatology, meteorology and information technology, provides an integrated approach to solving the problem of methane emissions. This emphasises the importance of researchers and practitioners working together to achieve effective solutions in environmental monitoring and climate risk management.

CONCLUSIONS

The study determined that the innovative RTM-based satellite data analysis and methane detection software significantly improved the accuracy of methane monitoring. This software, which uses radiative transfer algorithms, can be used to improve the efficiency of important physical processes analysis, such as the scattering and absorption of infrared radiation in the atmosphere. Considering the influence of atmospheric factors such as cloudiness and aerosols, a significant reduction in calculation errors was achieved, which is critical for assessing the impact of methane on the climate.

The analysis of spatial and temporal changes in methane concentrations has shown that methane emissions can be irregular and depend on many factors, such as agricultural activities, the oil and gas industry and natural phenomena. This highlights the need for accurate monitoring, which can serve as a basis for developing effective emission reduction strategies. As a result, the development of the Carbon Mapper system requires a comprehensive approach, including the selection of sensors, the creation of a reliable data collection infrastructure and the development of powerful analysis software. This will allow for effective monitoring of methane concentrations and contribute to the fight against climate change by providing scientific data for informed decision-making.

Testing the Carbon Mapper system is an important step in ensuring its performance and accuracy. Field trials, benchmarking and data processing with reporting will not only evaluate the system's effectiveness but also make an important contribution to methane monitoring, which in turn will help fight climate change. In addition, the use of modern data processing algorithms ensured high speed of analysis of large amounts of information coming from various satellite systems in real-time. This is critical for global monitoring, where efficiency is a key factor.

Overall, the results of the study confirm that an interdisciplinary approach based on the cooperation of specialists from various fields, such as climatology, meteorology and information technology, is central to the successful development of software for monitoring greenhouse gas emissions. This collaboration can incorporate many parameters and ensure high accuracy of the results, which is an important step towards a more effective fight against climate change. It is necessary to further study the impact of various anthropogenic and natural factors on the dynamics of methane emissions to improve the accuracy of forecasts and develop more effective strategies to reduce them. The limitation of the study is the possibility of the unavailability of some satellite data in real time, which may affect the accuracy and timeliness of the analysis of methane emissions.

ACKNOWLEDGEMENTS

None.

CONFLICT OF INTEREST

None.

REFERENCES

[1] Allen, D., *et al.* (2022). Microplastics and nanoplastics in the marine-atmosphere environment. *Nature Reviews Earth & Environment*, 3(6), 393-405. doi: 10.1038/s43017-022-00292-x.

[2] Asha, P., Natrayan, L., Geetha, B.T., Beulah, J.R., Sumathy, R., Varalakshmi, G., & Neelakandan, S. (2022). IoT enabled environmental toxicology for air pollution monitoring using AI techniques. *Environmental Research*, 205, article number 112574. doi: 10.1016/j.envres.2021.112574.

[3] Balasus, N., *et al.* (2023). A blended TROPOMI+GOSAT satellite data product for atmospheric methane using machine learning to correct retrieval biases. *Atmospheric Measurement Techniques*, 16(16), 3787-3807. doi: 10.5194/amt-16-3787-2023.

[4] Bruhwiler, L., Basu, S., Butler, J.H., Chatterjee, A., Dlugokencky, E., Kenney, M.A., McComiskey, A., Montzka, S.A., & Stanitski, D. (2021). Observations of greenhouse gases as climate indicators. *Climatic Change*, 165(1), article number 12. doi: 10.1007/s10584-021-03001-7.

[5] Bui, L.T., Nguyen, P.H., & Nguyen, D.C.M. (2021). A web based methane emissions modelling platform: Models and software development. *Waste Management*, 134, 120-135. doi: 10.1016/j.wasman.2021.08.015.

[6] Collins, W., Orbach, R., Bailey, M., Biraud, S., Coddington, I., DiCarlo, D., Peischl, J., Radhakrishnan, A., & Schimel, D. (2022). Monitoring methane emissions from oil and gas operations. *Optics Express*, 30(14), 24326-24351. doi: 10.1364/OE.464421.
[7] Cooper, J., Dubey, L., & Hawkes, A. (2022). Methane detection and quantification in the upstream oil and gas sector: The role of satellites in emissions detection, reconciling and reporting. *Environmental Science: Atmospheres*, 2(1), 9-23. doi: 10.1039/D1EA00046B.
[8] Douros, J., *et al.* (2023). Comparing Sentinel-5P TROPOMI NO_2 column observations with the CAMS regional air quality ensemble. *Geoscientific Model Development*, 16(2), 509-534. doi: 10.5194/gmd-16-509-2023.
[9] Erland, B.M., Thorpe, A.K., & Gamon, J.A. (2022). Recent advances toward transparent methane emissions monitoring: A review. *Environmental Science & Technology*, 56(23), 16567-16581. doi: 10.1021/acs.est.2c02136.
[10] Gålfalk, M., Nilsson Påledal, S., & Bastviken, D. (2021). Sensitive drone mapping of methane emissions without the need for supplementary ground-based measurements. *ACS Earth and Space Chemistry*, 5(10), 2668-2676. doi: 10.1021/acsearthspacechem.1c00106.
[11] Global Methane Tracker. (2024). Retrieved from https://www.iea.org/reports/global-methane-tracker-2024.
[12] Hui, D., Deng, Q., Tian, H., & Luo, Y. (2022). Global climate change and greenhouse gases emissions in terrestrial ecosystems. In M. Lackner, B. Sajjadi & W.-Y. Chen (Eds.), *Handbook of climate change mitigation and adaptation* (pp. 23-76). Cham: Springer. doi: 10.1007/978-3-030-72579-2_13.
[13] Imasu, R., *et al.* (2023). Greenhouse gases Observing SATellite 2 (GOSAT-2): Mission overview. *Progress in Earth and Planetary Science*, 10(1), article number 33. doi: 10.1186/s40645-023-00562-2.
[14] Jacob, D.J., *et al.* (2022). Quantifying methane emissions from the global scale down to point sources using satellite observations of atmospheric methane. *Atmospheric Chemistry and Physics*, 22(14), 9617-9646. doi: 10.5194/acp-22-9617-2022.
[15] Jamshed, W., Nisar, K.S., Gowda, R.J.P., Kumar, R.N., & Prasannakumara, B.C. (2021). Radiative heat transfer of second grade nanofluid flow past a porous flat surface: A single-phase mathematical model. *Physica Scripta*, 96(6), article number 064006. doi: 10.1088/1402-4896/abf57d.
[16] Jiang, Y., Zhang, L., Zhang, X., & Cao, X. (2024). Methane retrieval algorithms based on satellite: A review. *Atmosphere*, 15(4), article number 449. doi: 10.3390/atmos15040449.
[17] Johnson, M.S., Matthews, E., Du, J., Genovese, V., & Bastviken, D. (2022). Methane emission from global lakes: New spatiotemporal data and observation-driven modeling of methane dynamics indicates lower emissions. *Journal of Geophysical Research: Biogeosciences*, 127(7), article number e2022JG006793. doi: 10.1029/2022JG006793.
[18] Li, J., Tian, Y., Zhang, Y., & Xie, K. (2021). Spatializing environmental footprint by integrating geographic information system into life cycle assessment: A review and practice recommendations. *Journal of Cleaner Production*, 323, article number 129113. doi: 10.1016/j.jclepro.2021.129113.
[19] Liu, Q., Yan, B., Garrett, K., Ma, Y., Liang, X., Huang, J., Wang, W., & Cao, C. (2022a). Deriving surface reflectance from visible/near infrared and ultraviolet satellite observations through the Community Radiative Transfer Model. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15, 2004-2011. doi: 10.1109/JSTARS.2022.3149767.
[20] Liu, X., Wang, L., Kong, X., Ma, Z., Nie, B., Song, D., & Yang, T. (2022b). Role of pore irregularity in methane desorption capacity of coking coal. *Fuel*, 314, article number 123037. doi: 10.1016/j.fuel.2021.123037.
[21] Lovrak, A., Pukšec, T., Grozdek, M., & Duić, N. (2022). An integrated Geographical Information System (GIS) approach for assessing seasonal variation and spatial distribution of biogas potential from industrial residues and by-products. *Energy*, 239(B), article number 122016. doi: 10.1016/j.energy.2021.122016.
[22] Montes-Pérez, J.J., Obrador, B., Conejo-Orosa, T., Rodríguez, V., Marcé, R., Escot, C., Reyes, I., Rodríguez, J., & Moreno-Ostos, E. (2022). Spatio-temporal variability of carbon dioxide and methane emissions from a Mediterranean reservoir. *Limnetica*, 41(1), 43-60. doi: 10.23818/limn.41.04.
[23] Omara, M., *et al.* (2023). Developing a spatially explicit global oil and gas infrastructure database for characterizing methane emission sources at high resolution. *Earth System Science Data*, 15(8), 3761-3790. doi: 10.5194/essd-15-3761-2023.
[24] Palmer, P.I., Feng, L., Lunt, M.F., Parker, R.J., Bösch, H., Lan, X., Lorente, A., & Borsdorff, T. (2021). The added value of satellite observations of methane for understanding the contemporary methane budget. *Philosophical Transactions of the Royal Society A*, 379(2210), article number 20210106. doi: 10.1098/rsta.2021.0106.
[25] Shkundalov, D., & Vilutienė, T. (2021). Bibliometric analysis of building information modeling, geographic information systems and web environment integration. *Automation in Construction*, 128, article number 103757. doi: 10.1016/j.autcon.2021.103757.
[26] Stegmann, P.G., Johnson, B., Moradi, I., Karpowicz, B., & McCarty, W. (2022). A deep learning approach to fast radiative transfer. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 280, article number 108088. doi: 10.1016/j.jqsrt.2022.108088.

[27] Terrenoire, E., Hauglustaine, D.A., Cohen, Y., Cozic, A., Valorso, R., Lefèvre, F., & Matthes, S. (2022). Impact of present and future aircraft NO_x and aerosol emissions on atmospheric composition and associated direct radiative forcing of climate. *Atmospheric Chemistry and Physics*, 22(18), 11987-12023. doi: 10.5194/acp-22-11987-2022.
[28] Tyner, D.R., & Johnson, M.R. (2021). Where the methane is – insights from novel airborne LiDAR measurements combined with ground survey data. *Environmental Science & Technology*, 55(14), 9773-9783. doi: 10.1021/acs.est.1c01572.
[29] Wu, X., *et al.* (2024). Advances in methane emissions from agricultural sources: Part I. Accounting and mitigation. *Journal of Environmental Sciences*, 140, 279-291. doi: 10.1016/j.jes.2023.08.029.
[30] Zhang, S., Ma, J., Zhang, X., & Guo, C. (2023). Atmospheric remote sensing for anthropogenic methane emissions: Applications and research opportunities. *Science of the Total Environment*, 893, article number 164701. doi: 10.1016/j.scitotenv.2023.164701.

Інноваційне програмне забезпечення для аналізу супутникових даних і викидів метану із застосуванням моделі радіаційного переносу

Камала Агаєва

Доктор філософії, заступник директора з наукових питань
Азербайджанське національне аерокосмічне агентство
AZ 1115, вул. Ахундова, 1, м. Баку, Азербайджан
https://orcid.org/0009-0006-6779-2550

Герман Краукліт

Аспірант
Азербайджанське національне аерокосмічне агентство
AZ 1115, вул. Ахундова, 1, м. Баку, Азербайджан
https://orcid.org/0000-0002-6171-9348

Анотація. Дослідження було проведено для аналізу ефективності застосування моделі радіаційного переносу (RTM) в програмному забезпеченні для опрацювання супутникових даних і моніторингу викидів метану. У процесі дослідження було використано методи аналізу супутникових даних, моделювання радіаційного переносу та інтеграції з геоінформаційними системами для вивчення викидів метану та їхніх просторово-часових змін. Під час дослідження було встановлено, що застосування RTM для аналізу супутникових даних істотно підвищує точність оцінки викидів метану. Експериментальні дані засвідчили, що використання цієї моделі дає змогу більш ефективно враховувати атмосферні чинники, як-от хмарність та аерозолі, що мінімізує помилки в розрахунках концентрацій метану. Також було підтверджено можливість застосування цього підходу для моніторингу викидів у різних географічних регіонах із високою точністю. Супутникові дані дали змогу визначити ключові джерела метанових викидів, включно з промисловими зонами та природними джерелами. У результаті дослідження було виявлено, що програмне забезпечення Carbon Mapper може використовуватися як інструмент для глобального моніторингу метану та інших парникових газів, що сприяє більш ефективній боротьбі зі зміною клімату. Програмне рішення також інтегрується з геоінформаційними системами для надання візуалізації даних і поліпшення їхньої інтерпретації. Крім того, результати засвідчили, що RTM дає змогу точно визначати тимчасові зміни в концентраціях метану, що важливо для оперативного реагування на зростання викидів у критичних зонах. Програмне забезпечення продемонструвало високий ступінь масштабованості, що дає змогу застосовувати його для аналізу даних як локального, так і глобального масштабу. Таким чином, використання даної моделі в поєднанні з високоточним супутниковим моніторингом підтвердило свою ефективність в екологічному моніторингу та управлінні викидами парникових газів

Ключові слова: Carbon Mapper; хмарність; аерозолі; моніторинг; географічні регіони

ВІСНИК
Черкаського державного технологічного університету

Науковий збірник

Том 29, № 3. 2024

Заснований у 1996 р. Виходить чотири рази на рік

Оригінал-макет видання виготовлено у редакційно-видавничому відділі
Черкаського державного технологічного університету

Відповідальний редактор:
Я. Бурлака

Редагування англомовних текстів:
С. Воровський, К. Касьянов

Комп'ютерна верстка:
О. Глінченко

Підписано до друку 21 жовтня 2024 р.
Формат 60*84/8
Умов. друк. арк. 9,1
Наклад 50 прим.

Адреса видавництва:
Черкаський державний технологічний університет
18006, бульв. Шевченка, 460, м. Черкаси, Україна
Тел.: +380638013283
E-mail: info@bulletin-chstu.com.ua
https://bulletin-chstu.com.ua/uk

BULLETIN
of Cherkasy State Technological University

Scientific Collection

Volume 29, No. 3. 2024

Founded in 1996. Published four times per year

The original layout of the publication is made in the editorial and publishing department of Cherkasy State Technological University

Managing editor:
Ya. Burlaka

Editing English-language texts:
S. Vorovsky, K. Kasianov

Desktop publishing:
O. Glinchenko

Signed for print of October 21, 2024.
Format 60*84/8
Conventional printed pages 9.1
Circulation 50 copies

Editors Office Address:

Cherkasy State Technological University
18006, 460 Shevchenko Blvd., Cherkasy, Ukraine
Tel.: +380638013283
E-mail: info@bulletin-chstu.com.ua
https://bulletin-chstu.com.ua/en